Data Density and Trend Reversals in Auditory Graphs: Effects on Point Estimation and Trend Identification Tasks

A Thesis
Presented to
The Academic Faculty

by

Michael A. Nees

In Partial Fulfillment
of the Requirements for the Degree
Master of Science in the
School of Psychology

Georgia Institute of Technology
May 2007

Data Density and Trend Reversals in Auditory Graphs: Effects on Point Estimation and Trend Identification Tasks

Approved by:

Dr. Bruce N. Walker, Advisor
School of Psychology
Georgia Institute of Technology

Dr. Gregory Corso
School of Psychology
Georgia Institute of Technology

Dr. Wendy Rogers
School of Psychology
Georgia Institute of Technology

Date Approved: January 5, 2007

ACKNOWLEDGEMENTS

I want to thank my parents, Danny and Susan Nees; they have always found ways to ensure that opportunities were available to me. I would not be successful in graduate school without the constant support and encouragement of my parents, as well as Scott Nees and the Yost family. My wife, Anne Tucker Nees, has been my biggest and most unwavering supporter in this undertaking; I thank her for everything, including seeing me through this project. This thesis reflects the mentorship of Terri Bonebright, who guided me not only to the study of psychology but also to Georgia Tech. My advisor, Bruce Walker, has been incredibly supportive, perpetually available, and a tremendous positive influence on all of my graduate work, including this thesis, and I am both fortunate and grateful to have conducted this work with his help and guidance. Wendy Rogers and Greg Corso offered valuable insights that helped to mold my initial proposal into a greatly improved final product, and I learned a great deal from their helpful feedback. I thank our research assistant, Dianne Palladino, for her hard work during data collection as well as her interest in this project. Finally, I would like to thank the members of the Sonification Lab; their friendship and companionship have made graduate school a unique and enjoyable experience on a daily basis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

Auditory graphs—displays that represent graphical, quantitative information with sound—have the potential to make graphical representations of data more accessible to blind students and researchers as well as sighted people. No research to date, however, has systematically addressed the attributes of data that contribute to the complexity (the ease or difficulty of comprehension) of auditory graphs. A pair of studies examined the role of both data density (i.e., the number of discrete data points presented per second) and the number of trend reversals for both point estimation and trend identification tasks with auditory graphs. For the point estimation task, results showed main effects of both variables, with a larger effect attributable to performance decrements for graphs with more trend reversals. For the trend identification task, a large main effect was again observed for trend reversals, but an interaction suggested that the effect of the number of trend reversals was different across lower data densities (i.e., as density increased from 1 to 2 data points per second). Results are discussed in terms of data sonification applications and rhythmic theories of auditory pattern perception.

CHAPTER 1

INTRODUCTION

Graphical representations of quantitative information play an important role in the dissemination of information in both public and scientific forums. Jones and Carreras (1996) estimated that 2.2 trillion graphs were published via print media in 1994 alone. Zacks, Levy, Tversky, and Schiano (2002) sampled academic journals and newspapers across ten years (1985-1994). They found that the average number of graphs published per issue in academic journals rose from 34.7 to 61.2 during the sampled time period, while the average number of graphs appearing in newsprint rose from 10.1 to 24.5 per issue. Not surprisingly, Peden and Hausmann (2000) found an average of 67.73 graphs per book in a sample of introductory psychology texts.

Graphical information displays pervade published material, because they can offer a relatively concise visual summary of data. Furthermore, properly designed graphs can facilitate the emergence of data features (e.g., patterns, see Sanderson, Flach, Buttigieg, & Casey, 1989) that are not immediately evident in non-graphical data depictions; thus, researchers (e.g., Kosslyn, 1989) have theorized such data features in graphical displays are automatically perceived by the viewer. As a result of their prevalence and advantages over text for quantitative information display, graphs have been empirically studied and discussed both as instructional aids for learning in children or novice populations (e.g., Gobbo, 1994; Lee & Gerber, 1999; Liu, Salvendy, & Kuczek, 1999; Moore, 1993) and more generally as communicative information displays (see Butler, 1993; Carpenter & Shah, 1998; Carswell, 1992; Gillan & Lewis, 1994; Kosslyn, 1989; Meyer, Shinar, & Leiser, 1997). The representation of quantitative data, however, need not be confined to visual displays; scientists began to investigate the potential to represent data in the auditory modality over 50 years ago (for a brief review, see Frysinger, 2005). More recently, researchers have begun to examine auditory graphs as tools for data exploration and analysis.

1.1 Rationale: The Need for Auditory Graphs

The most obvious motivation for pursuing sound as a means of quantitative information display has been the lack of access to graphical visualizations for blind and visually impaired people. Kramer (1994), however, reviewed a number of users, tasks, and scenarios where auditory displays could be advantageous (e.g., during eyes-busy tasks, when line of sight is obscured, etc.), and researchers have argued that auditory graphs have the potential to facilitate comprehension of graphical information for both blind and sighted students and scientists (Flowers, Buhman, & Turnage, 2005; Walker & Nees, 2005a).

The World Health Organization estimated that there are 161 million visually impaired people worldwide, 4.7 million of whom are in the United States (Resnikoff et al., 2004). Advances in computing technology and the advent of the world wide web have spurred the widespread publication of information in electronic form, and screen readers like Jaws (www.freedomscientific.com) coupled with a push toward accessible web design (e.g., Mankoff, Fait, & Tran, 2005; Takagi, Asakawa, Fukuda, & Maeda, 2004) have opened a tremendous amount of previously unavailable information to blind computer users. A deficiency exists, however, in the ability of screen readers to represent graphical material. For blind or visually impaired users of software and the world wide web, screen readers display information through a text-to-speech conversion. Images may be “tagged” with a hidden text description of the image, such as when a screen reader encounters a picture of a dog and reads aloud the tag “picture of dog.” This commonly used and often recommended (e.g., Massachusetts Institute of Technology, n.d.; Quesenbery, 1999; W3C, 2000) tagging strategy may make some media (i.e., pictures, photographs, etc.) suitably accessible for the blind. Graphs like those that pervade print media, however, are not easy to represent with a simple text tag. Furthermore, any attempt to verbally describe a graphical display of quantitative information results in a lengthy text description of the graph that may remove the advantages of the graphical display. To make the quantitative information “accessible,” tables have often been used to represent data in lieu of more appropriate line graphs (for some guidelines for presenting information in graphical form, see American Psychological Association, 2001; Gillan, Wickens, Hollands, & Carswell, 1998; Oliver, 1998), a practice that also may remove the advantages of graphical representations. Other current practices in graphing for the visually impaired include the use of tactile or Braille graphs, but the hardware to produce tangible graphics has been cost prohibitive.

Regarding the potential for auditory displays to prove useful for sighted persons, a number of scenarios have been proposed whereby auditory displays might be advantageous over traditional visual methods of information display. Some environmental conditions or constraints of a system render visual displays unusable (such as when a person’s line of sight with the display is obscured, Kramer, 1994) or insufficient (like in cellular phones or other mobile devices where screen sizes are exceedingly small, see Brewster, 2002). On the other hand, technological and computing advances have also allowed for substantial *increases* in the size and number of visual displays in the typical workstation, and this extra screen space can introduce problems, particularly with the overabundance and disorganization of visual information (for a discussion, see Grudin, 2001). Indeed, Wickens' (e.g. , 2002) description of multiple resources makes theoretical predictions that suggest less dual task interference when information can be spread across modalities rather than processed within a single modality (e.g., by vision alone). Studies have confirmed that auditory displays can be beneficial in scenarios where the display is small (Brewster, 2002; Brewster & Murray, 2000) or when the display is such that vision may be overtaxed (Brewster, 1997; M. L. Brown, Newsome, & Glinert, 1989).

Coincident with advances in visual computing displays, digital technology has allowed for the economical and widespread production and implementation of sounds via personal computers (Flowers et al., 2005). Auditory representations of quantitative information could fill a gap in data accessibility for the blind or help sighted people interact with computers and other digital technologies more efficiently. Technological advances in sound production capabilities combined with sparse literature and theory on sonification, however, have resulted in an exceptionally large design space for auditory graph builders. For example, Levitin (1999) proposed seven attributes of sound that can be manipulated; thus, data could be represented by sound along a multitude of different dimensions.

1.2 What is an Auditory Graph?

The theoretical foundation of precisely how to use sound to represent graphical relationships seems to lie in Kubovy's Theory of Indispensable Attributes (TIA, see Kubovy, 1981, 1988; van Valkenburg & Kubovy, 2003), although auditory graph literature has not made this connection explicit. Kubovy's theory proposed that the two indispensable attributes of vision (i.e., those stimulus properties that *must* be varied to produce the perception of distinct entities or objects) are space and time. Kubovy (1981) has further proposed that frequency and time are the two indispensable attributes of audition, which lead to the conclusion that "visual spatial location is analogous to auditory frequency" (pp. 78). Not surprisingly, then, auditory graphs most commonly map data changes along the visual Y-axis to changes in the frequency of sounds, while visual X-axis changes are mapped to the temporal presentation of the sound (see, for example, Bonebright, Nees, Connerley, & McCain, 2001; L. M. Brown & Brewster, 2003; L. M. Brown, Brewster, Ramloll, Burton, & Riedel, 2003; Childs, 2005; Flowers & Hauer, 1995; Mansur, Blattner, & Joy, 1985; Smith & Walker, 2002; Smith & Walker, 2005; Walker & Nees, 2005b).

1.3 Empirical Investigations: Can Listeners Use Auditory Graphs?

An early study of auditory graphs found them to be comparable in efficacy to the tactile displays traditionally used to present quantitative information to the blind (Mansur et al., 1985). Tactile data representations yielded slightly more accurate responses, while auditory graphs resulted in faster reaction times. Flowers and Hauer (1993) found that auditory graphical representations were better than visual representations of the same data for imparting information regarding central tendency and range. They later demonstrated that participants categorize sets of visual graphs and their auditory counterparts as perceptually similar or dissimilar along the same dimensions, namely shape, slope, and degree of linearity (1995).

Another study by Flowers, Buhman, and Turnage (1997) gave further evidence for the comparability of auditory and visual presentations of scatterplots. Participants estimated the same Pearson correlation *r* value for visual scatterplots and their corresponding auditory graphs. More recently, Bonebright et al. (2001) determined that, in general, participants were able to match an auditory graph to a visual line graph or scatterplot of the same data, and Brown and Brewster (2003) found that people could produce a visual rendition of a graph that was over 80% accurate (on average) after hearing a sonified presentation.

1.4 Insights from Visual Graph Comprehension Literature on Graph Complexity

Despite the promising potential of auditory graphs as a viable information display, studies of the data properties that contribute to auditory graph complexity are lacking. The extensive literature on visual graph comprehension, however, can potentially offer insight into those factors affecting graph complexity. Several models and theories of graph comprehension exist (e.g., Carpenter & Shah, 1998; Carswell, 1992; Gillan & Lewis, 1994; Kosslyn, 1989; Pinker, 1990). Such theories and models cannot necessarily be directly applied to auditory graphs, and auditory graph researchers cannot assume that auditory graph comprehension is a simple translation of visual graph processes to another modality. Nevertheless, auditory graph researchers would be remiss to ignore the extensive literature and theory on visual graph comprehension, and the existence of parallels between auditory and visual graph comprehension processes seems a reasonable proposition if frequency in audition is analogous to space in vision (see Kubovy, 1981).

An early study of graph comprehension found that performance on a trend identification task decreased as the number of data points portrayed in a line graph increased (Schutz, 1961), which suggested that the number of data points in a graph contribute to complexity. Carswell (1993; Carswell & Ramzy, 1997) has operationally defined graphical complexity as a function of the graph's number of data points, symmetry, and number of trend reversals, whereby more data points, asymmetrical functions, and more trend reversals were factors associated with greater graphical complexity. Meyer (2000; Meyer et al., 1997) defined graphical complexity according to the number of data points in the graph (more data points being more complex), the number of data series concurrently presented (where increasing the number of series is indicative of greater complexity), and the graphed data's pattern (e.g., the presence of meaningful data trends reduces complexity). Similarly, Friel, Curcio, and Bright (2001) claimed that graphical complexity "refers not only to the number of data items or categories but also to the kinds of data types (e.g., discrete vs. continuous), the spread and variation within the data set, and so on" (Progression of Graphs for Instruction section, para. 3). Clearly, as Friel et al.'s "and so on" conclusion elucidates, even in the well-established literature of visual graph comprehension, descriptions of the factors contributing to the complexity of graphs vary, although researchers seem to agree that the number of data points, the number of data series, and the trend and shape of data patterns should impact graph comprehension.

1.5 Data Complexity in Auditory Graph Literature

In contrast to the various descriptions of factors contributing to data complexity in the visual graph comprehension literature, auditory graph studies to date have not manipulated fundamental characteristics of sonified data to investigate the data properties that contribute to complexity. Previous research, however, has suggested some data characteristics that might contribute to difficulty of comprehension. Consistent with suggestions from visual graph literature, Bonebright et al. (2001) found that auditory graphs portraying two data series (where data series were separated spatially such that one data series was presented to each ear) were more difficult and took longer to match to their visual counterpart. They also found that more disperse scatterplots were more difficult to accurately match to their visual representations.

Roth, Kamel, Petrucci, and Pun (2002) found that participants had great difficulty (only a 25% rate of success) at identifying an auditory line graph as linear increasing; most participants mistook this function for a parabola. Performance for identifying auditory representations of parabola and sine wave functions was better (both had 75% rates of success). Roth et al.'s curious findings regarding poor performance for the identification of a simple linear increasing function, however, may be an artifact of their sonification technique. They mapped Y axis values not only to frequency changes, but also to changes in spatial elevation in the headphones. Although such redundant mappings are generally believed to be helpful in sonification (Kramer, 1994), auditory spatial elevation judgments are known to be susceptible to inaccuracies (see, for example, Wenzel, Arruda, Kistler, & Wightman, 1993). It seems possible that the unusual choice of a secondary spatial mapping led to the detriment in performance observed for simpler functions, as no theory of graph or auditory pattern comprehension would predict this pattern of results for a frequency mapping alone. Furthermore, previous auditory graph research had found that participants could distinguish linear from exponential functions with 84% accuracy (Mansur et al., 1985).

The studies mentioned above offer insight into properties of data that might contribute to graph comprehension difficulty, but the majority of published auditory graph studies focus on comparisons across modalities or evaluations of different methods of sonification. No investigations that systematically manipulate data properties have been performed, and meta-analytic comparisons of different properties of sonified data sets that contribute to complexity (such as the number of data points, trend reversals, or data series displayed) are not possible across studies.

1.6 Insights and Theoretical Considerations from Auditory Perception Literature

While the literature specific to auditory graphs is sparse, stimuli like auditory graphs have been examined in other contexts in psychology and related disciplines, such as information theory and auditory pattern perception. Although the study of the perception of frequency in time is not new, sonification and auditory graph studies have required listeners to perform novel *tasks* (e.g., data exploration, monitoring, etc., see Kramer, 1994) with sounds as compared to past research on basic auditory perception. Although much of the sonification research has proceeded with little overarching theoretical guidance (Pirhonen, Murphy, McAllister, & Yu, 2006), existing theoretical approaches may provide valuable insights and make relevant predictions regarding the perception of auditory graphs, despite the unique tasks required of auditory graph listeners.

Early information theorists, for instance, took a mathematical approach to the transmission of periodic data (see Nyquist, 1928; Shannon, 1998/1949), including but not limited to the propagation of sound waves in time. Specifically, Nyquist showed as the frequency of any waveform (not necessarily sound) increases (i.e., as the waveform exhibits more changes in direction per unit of time), more discrete pieces of information about the waveform must be sampled per unit of time to define the waveform to an arbitrarily acceptable degree of precision. Although this approach was concerned with the mathematical limits of information transmission (and not the psychological repercussions of information transmission on a human perceiver), it has yet to be seen whether Nyquist's theory can be applied to predict adequate sampling rates of discrete tones per unit of time in the presence of trend reversals in auditory graphs.

Regarding psychological theories that might predict auditory graph performance, seminal research in attending to auditory information has focused mainly on competing sources of sound and the ability to monitor multiple sources and switch between sources (e.g., selective attention, for a review see Moray, 1969). Later theories (Deutsch & Feroe, 1981; M. R. Jones, 1976), however, were concerned with how people attend to and process sequential or serial stimuli *within* a single sound source. Although auditory graph research has ventured into the realm of multiple concurrently presented data streams (see Bonebright et al., 2001; L. M. Brown, Brewster, & Riedel, 2002), the ways in which basic changes within a single data stream affect auditory graph comprehension across tasks have not been fully explored. The current studies, then, examined only single data series graphs and drew insight from the literature regarding processing *within* a single stream of auditory stimuli; traditional selective attention approaches to processing auditory stimuli, however, would likely be highly relevant for predicting performance with auditory graphs presenting multiple concurrent data series.

M.R. Jones' (1976) rhythmic theory drew upon the earlier propositions of the rhythmic theory of behavior proposed by Martin (1972). Martin suggested that rhythm—

the “temporal patterning” (p. 487) of events in the environment—was central to information processing. He proposed that many serial stimuli (like speech, music, sequences of tones, etc.) are structured in patterns in time that can be described in terms of a hierarchy of relative time and relative accents, which may involve non-temporal aspects of the stimulus. Martin posited a role for the cyclical focusing of attention in anticipation of rhythmic events, and he suggested that certain stimulus configurations—namely those that propagate rhythms that follow a listener’s expectancies within the context of the hierarchy—were more amenable to processing than others.

M.R. Jones (1976) greatly expanded upon the work of Martin (1972) by elaborating a rhythmic theory that was derived from studies of auditory perception. The relevant dimensions of Jones’ theory were frequency, loudness, and time, and she emphasized the interactive nature of relational changes in the frequency and loudness dimensions of auditory stimuli as they occur in time. The categorization of stimuli was again hierarchical, meaning that the range of temporal patterns from simple to complex can be broken down and described in terms of constituent smaller sequences that form meaningful sub-patterns. Rhythmic theory, like Martin’s theory, posited a role for a periodic attentional focus (later dubbed “entrainment,” see McAuley & Jones, 2003) that could synchronize with rhythmic stimuli and facilitate perception.

M.R. Jones (1976) also theorized that certain stimulus configurations would be more favorable to processing than others, and the relative ease of comprehension of the patterns was described along a continuum ranging from nominal to interval perceptual relations. For sounds that are nominally related, the listener can perceive only that the sounds are the same or different. Ordinal relations allow for the perception of one sound as having higher or lower frequency than its comparator. Finally, interval relations are perceived as having both a direction and a magnitude. While objectively all sound relations could be described in interval terms (i.e., direction and magnitude of frequency differences between sounds can always be objectively quantified), rhythmic theory described the subjective, perceptual experience of listeners as a function of the interactions of dimensions, of which frequency and time are most relevant to the current discussion.

M.R. Jones (1976) offered predictions regarding not only the complexity of auditory patterns, but also the theoretical conditions under which pattern perception breaks down entirely (i.e., the percept does not match the veridical stimulus). Simple patterns—those patterns that are most amenable to processing interval relationships (i.e., both the direction and magnitude of frequency changes)—were theorized to be those that featured regular interval changes in a single direction over time (i.e., frequencies increasing or decreasing at regular intervals). Sequences of tones with contour changes and irregular interval changes were more complex in the theory, and perhaps only amenable to perceptions of ordinal relations in extreme circumstances. Finally, in cases with very large interval jumps (i.e., large changes in frequency from tone to tone) and multiple changes in direction, the perception of a temporal sequence may collapse altogether. In this instance, the listener may report two streams of tones grouped as high and low in frequency rather than a perception of a single stream of alternating tones (also see Bregman, 1990), and in extreme cases only nominal relations between the sounds will be perceived.

While Martin's (1972) theory suggested that increases in the rate of tones—holding other dimensions constant—would not affect perception of the pattern (due to constant *relationships* of frequency change per *unit* of time), M.R. Jones (1976) set a theoretical upper limit on the amount of permissible frequency change between tones that would preserve the perception of the sequence across changes in time. In other words, at high rates of presentation, greater frequency intervals that were acceptably perceivable at slower rates may need to be changed to smaller frequency intervals for the sequence to be perceived accurately in time. M.R. Jones' theory suggests a continuum leading up to the upper limit of frequency change, whereby greater and irregular frequency changes in an auditory pattern are more difficult to process than unidirectional, regular frequency changes.

Deutsch and Feroe (1981) proposed a model of the perception of tone sequences that was very similar in concept to M.R. Jones' (1976) theory. Deutsch and Feroe suggested, like Jones, that tone sequences are represented hierarchically, with lower nodes breaking down the overall pattern into meaningful subsets. The initial grouping of sequences was proposed to be determined by Gestalt principles like proximity (in either frequency or time) and good continuation (such as when a series of sounds consistently increase in frequency), and the model predicts that patterns exhibiting better Gestalt groupings are easier to perceive. Deutsch and Feroe also suggested that large jumps in frequency intervals from tone to tone are detrimental to grouping. Expectancies again play a role, as listeners are theorized to anticipate tonal entities based on previous tones and primitive grouping principles. Tones that match anticipated patterns strengthen the representation of the sequence, and mismatches serve to elaborate the representation, often at higher levels in the hierarchy provided a hierarchical rhythmic structure is present.

Despite the fairly straightforward predictions of these theories (i.e., Deutsch & Feroe, 1981; M. R. Jones, 1976; Martin, 1972) regarding ease and difficulty in auditory pattern perception, these theories did not conceive of auditory patterns as graphs, nor were they intended to address the exact types of *tasks* (e.g., point estimation, global and local trend identification) that are required of auditory graph listeners. As a result, it is unclear the extent to which the above theoretical predictions regarding auditory pattern perception will generalize to predict stimulus complexity and performance in auditory graphs.

1.7 The Current Research

The systematic examination of the fundamental data properties that affect auditory graph comprehension should provide valuable insights for sonification researchers. Theories of auditory pattern perception (e.g., Deutsch & Feroe, 1981; M. R. Jones, 1976; Martin, 1972) make predictions regarding the perceptibility (i.e., complexity) of sequences of sounds, and these insights might extrapolate to the types of stimuli and tasks found in sonification literature. Such existing theories of auditory pattern perception may be useful for framing sonification research questions and predicting performance with auditory displays. Kosslyn (1989) suggested that visual graphs provide for the automatic perception of some data features, and certain data features (e.g., a simple linear increasing trend) likewise may be readily perceivable from an auditory graph. Similarly, more complex data features may make information extraction more difficult. The identification of those properties of the data that may affect auditory graph comprehension should permit auditory graph designers and listeners to better understand and predict the attributes of data (e.g., patterns, etc.) that will be easily perceived from an auditory graph. The identification of the characteristics of more complex or difficult data sets will also allow auditory graph researchers to demonstrate skill acquisition (i.e., a progression from mastery of simple to complex graphs) in longitudinal studies of training, as researchers will know the factors that contribute to "easy" versus "difficult" auditory graphs.

A pair of studies manipulated two fundamental properties of auditory graph stimuli—the density of data points and number of trend reversals displayed—to examine the effects of graph complexity for both point estimation (Experiment 1) and trend identification (Experiment 2) tasks. Although many different tasks can be required of graphs users, the identification of trends and the estimation of a Y axis value for a given X axis value are common activities that are representative of some of the typical graph user's basic information needs. Other data attributes such as symmetry and the initial direction of change were held constant or counterbalanced in these studies.

Of note, these studies examined auditory line graphs, where no more than one Y axis value was displayed for a given X axis value. Although auditory graph researchers have investigated auditory scatterplots (e.g., Bonebright et al., 2001; Flowers et al., 1997), box-whisker plots (Flowers & Hauer, 1992; Peres & Lane, 2003; Peres & Lane, 2005), histograms (Flowers & Hauer, 1993), and tabular data (Stockman, Hind, & Frauenberger, 2005), the majority of auditory graph research has examined auditory line graphs (e.g., Bonebright et al., 2001; Brewster & Murray, 2000; L. M. Brown et al., 2002; L. M. Brown & Brewster, 2003; Flowers & Hauer, 1995; Mansur et al., 1985; Roth et al., 2002; Smith & Walker, 2002, 2005; Turnage, Bonebright, Buhman, & Flowers, 1996; Walker & Nees, 2005b). This is not surprising, considering that line graphs accounted for 72.5% of graphs appearing in academic journals and 50.1% of all graphs appearing in newspapers in the sample obtained by Zacks et al. (2002). Likewise, Peden and Hausmann (2000) found that, of the average 67.73 graphs appearing in introductory psychology textbooks, 42.36 (63%) were line graphs.

1.8 Hypotheses

1.8.1 Hypothesis 1a: Point estimation and data density

For the point estimation task (Experiment 1), performance was predicted to decline as data density increased (i.e., as more data points were presented). This prediction was consistent with empirical findings and theoretical predictions from visual graph comprehension literature (e.g., Carswell et al., 1993; Carswell & Ramzy, 1997; Friel et al., 2001; Meyer, 2000; Meyer et al., 1997; Schutz, 1961) and makes intuitive sense, because with more data points the listener must select the target from a larger array of discrete tones. This prediction is somewhat in conflict with the predictions of rhythmic theories like those of Martin (1972), which predicted no perceptual decrements no matter the density of tones per second in auditory patterns, and of M.R. Jones, (1976), which predicted a decrement under circumstances of fast rates of presentation coupled with extreme frequency jumps (rather than as a function of the rate of presentation per se). Neither rhythmic theory, however, was intended to predict performance for this particular task (point estimation) with auditory patterns.

1.8.2 Hypothesis 1b: Point estimation and trend reversals

Performance on the point estimation task was also predicted to decline as the number of trend reversals in the displayed data increased. This prediction was consistent with empirical findings and theoretical predictions from visual graph comprehension literature (e.g., Carswell et al., 1993; Carswell & Ramzy, 1997). This hypothesis was also consistent with predictions regarding auditory pattern complexity proposed by the theories of M.R. Jones (1976) and Deutsch and Feroe (1981). I remained agnostic regarding the potential interaction of data density and trend reversals for point estimation. M.R. Jones' theory would certainly predict an interaction of these variables if the theoretical upper limit for frequency change per unit of time was exceeded, but her proposed limits were not violated by the stimuli in the current study.

1.8.3 Hypothesis 2a: Trend identification and trend reversals

Regarding trend identification with auditory graphs (Experiment 2), performance on trend identification was predicted to decrease as the number of trend reversals increased. This prediction was consistent with research in the visual graph comprehension literature that has posited a role for trend reversals in the complexity of graphical displays (e.g., Carswell et al., 1993; Carswell & Ramzy, 1997). This prediction was also consistent with predictions regarding auditory pattern complexity proposed by the theories of M.R. Jones (1976) and Deutsch and Feroe (1981), which both suggested that directional changes in frequency create more elaborate, complicated auditory patterns than stimuli with unidirectional pitch.

1.8.4 Hypothesis 2b: Trend identification interaction of density and trend reversals

For the trend identification task, an interaction was also predicted such that lower data density (i.e., fewer data points presented per second in the display) would result in worse trend identification performance as the number of trend reversals increased. In the theories of M.R. Jones (1976) and Deutsch and Feroe (1981), it is unclear whether this interaction should be present for the tonal stimuli and tasks of the current study, but Jones generally suggested that the perception of sequences of tones is more difficult when frequency changes are greater or more frequent per unit of time. Information theory, however, would suggest that when data are very sparse within auditory graphs, the extraction of trend information should prove more difficult than with high data density, as listeners have less information with which to identify trends in the data values across time (i.e., there are fewer tones within the interval to confirm or disconfirm the direction of frequency change). Lower sampling rates of discrete data points (i.e., lower density), then, should have a greater negative impact on trend identification for auditory graphs with more trend reversals (see Nyquist, 1928; Shannon, 1998/1949).

CHAPTER 2

EXPERIMENT 1: POINT ESTIMATION

Experiment 1 examined the effects of data density and the number of trend reversals within the sonified data on performance of a point estimation task with auditory graphs.

2.1 Method

2.1.1 Participants

Participants (*N* = 32; 16 males and 16 females) were recruited from undergraduate psychology courses at the Georgia Institute of Technology. Participants' mean age was 19.13 years (*SD* = 1.62), and the ages sampled ranged from 18 to 25 years old. Participants reported having played a musical instrument for an average of 3.56 (*SD* = 3.98) years, with 9 participants having never played an instrument and 11 having played an instrument for 5 years or more. Participants reported a mean of 2.94 (*SD* = 3.52) years of formal musical training (i.e., private or class instruction), and a mean of 3.19 (*SD* = 3.96) years of experience with reading musical notation. Participants had a mean of 0.28 (*SD* = 0.73) years of experience with stock trading or closely following the stock market. They also reported having taken a mean of 2.88 (*SD* = 2.47) college or advanced placement level business or economics courses.

2.1.2 Apparatus

Visual presentations (such as instructions and text presentations of questions during trials) were made on a 17 in (43.2 cm) Apple LCD computer monitor. Auditory presentations were delivered via Sennheiser HD 202 headphones, which were adjusted by the participant to a comfortable fit. Listening volume was approximately 65 dB SPL. All presentations of stimuli and data collection were accomplished with the Macromedia Director MX 2004 software package.

2.1.3 Data Sets for Stimuli

Auditory graphs depicted the price of a stock in dollars as it varied (within a range of 6 dollars to 106 dollars) over the course of a 10-hour trading day that opened at 8 a.m. and closed at 6 p.m. All auditory graphs were 10.125 seconds long. Fictional stock price data have been used in past research (e.g., Smith & Walker, 2005; Walker & Nees, 2005b), because they represent a relatively generic domain that should be accessible to naïve subjects with no specialized expertise in the area. To further ensure that domain knowledge did not influence performance for tasks in the current study, participants were given a brief overview of the task with information regarding the domain (i.e., stocks have monetary values that fluctuate during a day of market trading, etc.). No additional prior knowledge of the domain was necessary for the graphing tasks in the current study.

The price of the stock in dollars (on the Y-axis) was represented by discrete tones that changed in frequency as the price changed, while each hour of the trading day (on the visual X-axis) was mapped to one second in time. The duration of individual tones was held constant at 125 ms per tone, with 10 ms onset and offset ramps. The experimental manipulations of the current study involved 4 different levels of auditory graph data density and 4 different levels of trend reversals for a total of 16 different combinations of stimuli.

2.1.3.1 Data Density

Data density was manipulated at 4 levels that offered a psychologically and practically relevant range of stimuli: 1 data point per second, 2 data points per second, 4 data points per second, and 8 data points per second. Stimulus interonset intervals (IOIs) under 1800 ms promote perceptual grouping; with longer intervals items may be perceived independently rather than as members of a sequence (see Fraisse, 1978; Fraisse, 1982). Perception of tones as a coherent auditory graph, therefore, may fail if data are not presented at a minimally sufficient rate, and the lowest data density employed here falls well within the limit of grouping by IOI for tones. Furthermore, from a practical perspective, auditory graphs will need to be designed such that a listener can explore the data in a reasonable amount of time. For example, rhythmic sequences characterized by presentation rates around 1 item per second have been characterized as perceptually "slow", while about 5-6 items per second are perceived as "fast" (Palomaki, 2006), and in studies of tempo, researchers have operationally defined stimuli with at or near 4 items per second as "fast" (M. R. Jones, Johnston, & Puente, 2006) and at or near 10 items per second as "very fast" (Drake & Botte, 1993).

Of further interest regarding reasonable upper limits for the temporal presentation of data in auditory graphs, the threshold for determining the order of temporally presented stimuli (i.e., being able to perceive which item preceded adjacent items in a series) has been shown to range from 20 to 100 ms (Fraisse, 1978), and an auditory graph listener would be served poorly by a graph where this "threshold of succession" was ambiguous. The current study's fastest rate of presentation, while falling within a presentation rate that is perceived as "fast," was well below the rate whereby succession of items was indeterminate in any previous research. Another important consideration involved the duration of discrete, individual tones, which necessarily decreased as more tones were added per unit of time. Early research (Turnbull, 1944) suggested that frequency discrimination deteriorated rapidly as the length of a tone fell below durations of around 100 ms. The current study used tones of a duration well-above the threshold for failures of frequency discrimination due to tonal duration. Finally, using magnitude estimation techniques, Walker (2002) found that perception of the magnitude of tempo (as tempo) increased approximately linearly as tempo increases, thus each stimulus level for data density in the current study represented an approximate doubling of perceived tempo to cover the range from 1 to 8 tones per second.

2.1.3.2 Trend Reversals

The second experimental variable involved the number of trend reversals presented in the auditory graph stimuli. The 4 levels of the trend reversals variable were operationally defined as 0, 1, 2, or 3 trend reversals in the data. Auditory graphs with zero trend reversals represented data that either increased or decreased monotonically across the entire trading day. Graphs with one trend reversal data rose for the first half of the trading day then fell, or vice versa. Graphs with 2 and 3 trend reversals assumed 2 and 3 changes in the direction of the stock price data trend, respectively. As predicted in rhythmic theory (M. R. Jones, 1976), the number of changes in direction for frequency influences the perceptual complexity of an auditory pattern, with simplest patterns characterized by monotonic frequency changes. Similarly, Deutsch and Feroe (1981) called for the application of Gestalt principles to describe the complexity of tonal sequences, with unidirectional frequency changes representing the simplest patterns due to frequency proximity, temporal proximity, and good continuation of adjacent sounds. Complexity of auditory patterns, then, should increase as the number of frequency changes (i.e., trend reversals in auditory graphs) increase. Although in some domains real data may change more sporadically than the simple patterns employed for these stimuli, the levels for the trend reversals manipulation represented a starting point for exploring how changes in trends impact auditory graph performance.

Figures 1 and 2 offer visual depictions of each of the combinations of data density and trend reversals. Of note, for each factorial combination of the two independent variables, auditory graphs could be constructed such that the data in the graphs initially increase (Figure 1) or initially decrease (Figure 2). For a given factorial combination of data density and trend reversals from the 16 different stimulus combinations, participants experienced only trials with trends that initially increased or only trials with trends that initially decreased within a stimulus combination cell. Across stimulus conditions, a given participant experienced 50% of trials with initially increasing graphs and 50% of trials with initially decreasing graphs. For example, for data density of 1 data point per second coupled with zero trend reversals, half of the participants experienced linear increasing graphs, and the other half experienced linear decreasing graphs. The initial direction (increasing or decreasing) of the data was assumed to have no impact on participants' comprehension of the graph, but the initial direction for each of the 16 combinations of the independent variables was counterbalanced across participants to eliminate any potential unforeseen confounding with the variables of interest.

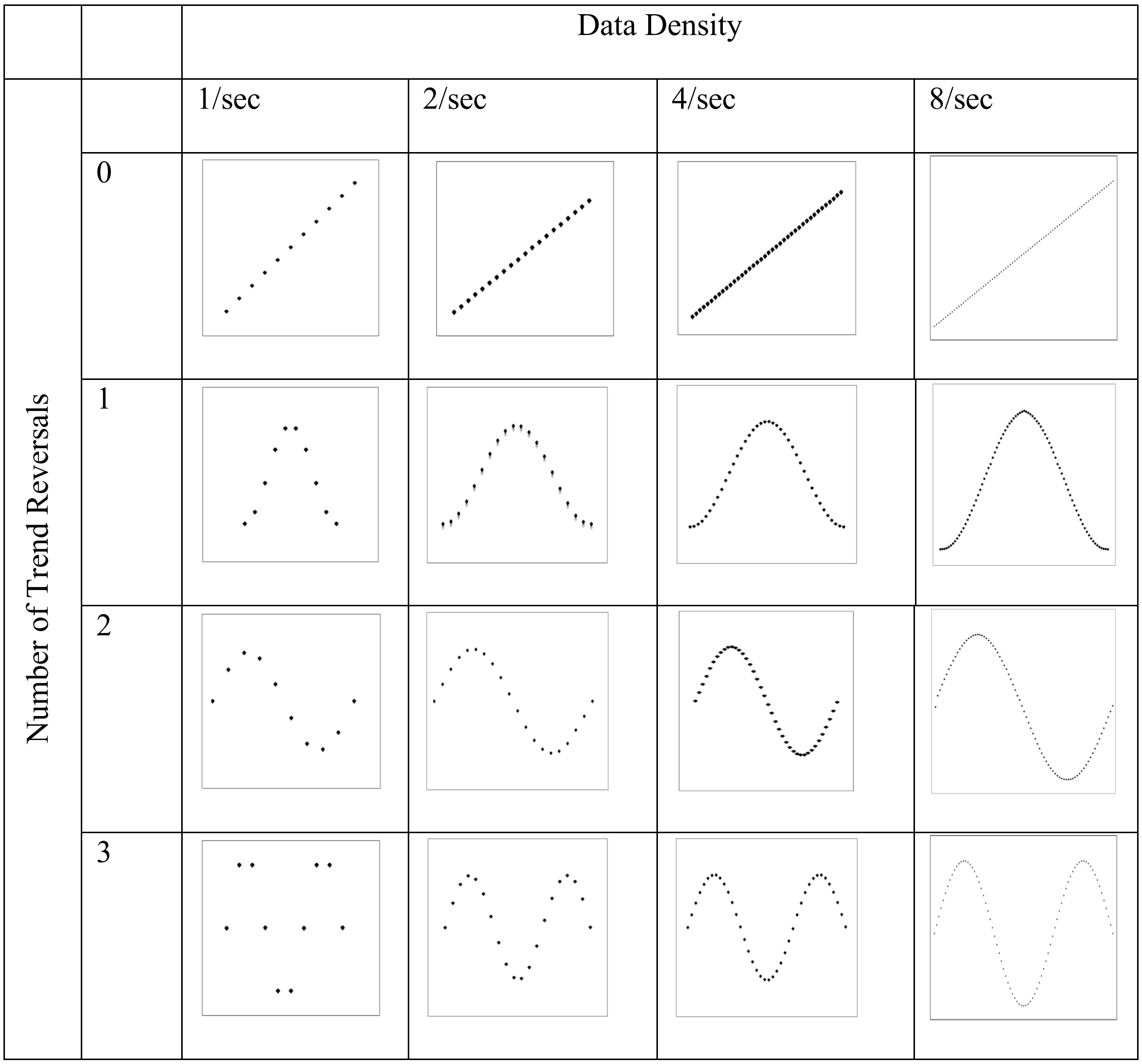

Figure 1. Visual depictions of the 16 factorial combinations of stimuli with initially increasing data values.

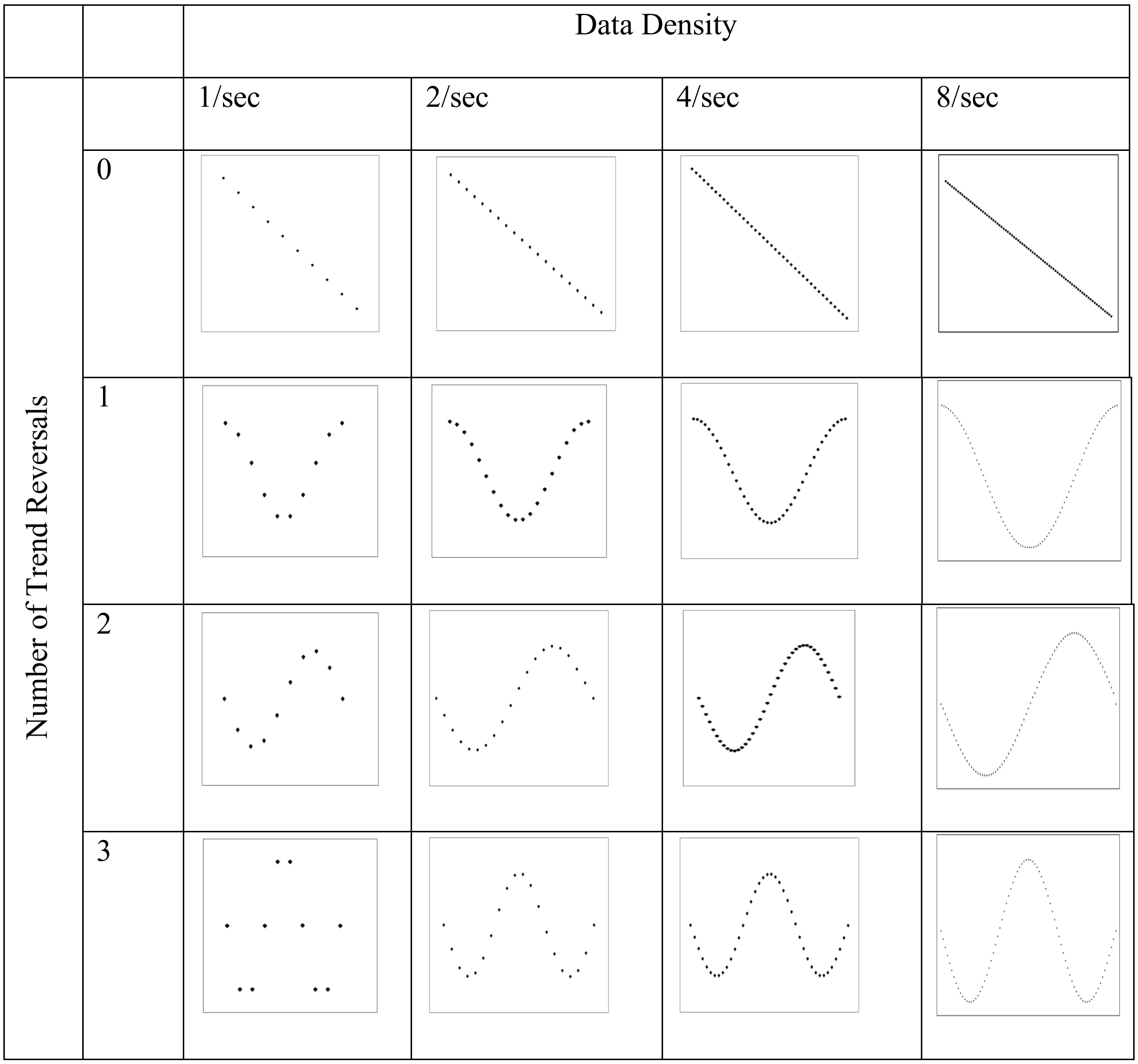

Figure 2. Visual depictions of the 16 factorial combinations of stimuli with initially decreasing data values.

2.1.4 Auditory Graph Stimuli Design

2.1.4.1 Sonification of Stock Price Data

Data for auditory graphs were sonified as MIDI data files using the *Sonification Sandbox* program (see Walker & Cothran, 2003; Walker & Lowey, 2004), converted to .WAV files, and exported to Audacity version 1.3.0b for track mixing. All sounds presented the same signal to both left and right headphone channels (i.e., a center mix without stereo panning was employed).

Within the *Sonification Sandbox* program, stock data were represented with the MIDI instrument bank's piano timbre. Scaling anchors were assigned for maximum and minimum values in a data set. The minimum data value ($6) was assigned MIDI note G2, whose frequency was 98 Hz, and the maximum data value ($106) corresponded to MIDI note B6, whose frequency was 1979.5 Hz. A positive polarity mapping was employed, and data values in between the maximum and minimum were assigned MIDI notes on an exact scale (i.e., in the event that data values fell between whole notes, the tones were adjusted to represent the exact frequency of the data point on the scale).

2.1.4.2 Auditory context for X and Y axes

Auditory context in the form of Y axis reference tones and X axis click tracks has been shown to generally provide cues that aid in the performance of point estimation tasks with auditory graphs (Smith & Walker, 2002, 2005). Despite the potential for auditory context to aid auditory graph comprehension, both Y axis reference tones and X axis click tracks were omitted in this study, as concurrent auditory context represented a potential confound when implemented with the manipulations of the current studies.

Regarding Y axis reference tones, the most useful reference tones alternate (from maximum to minimum references) as the data trends change (Smith & Walker, 2002, 2005). Given that the current study investigated the impact of trend reversals, the number of reference tone alternations across stimuli would change as a function of trend reversals, perhaps producing differential effects with regard to auditory stream segregation (Bregman, 1990) and contextual benefit.

Furthermore, considerable research has indicated that rhythmic perception is hierarchical, with lowest nodes at the level of individual discrete sounds, which are grouped according to accents (i.e., rhythmic beats), etc., proceeding upward to higher-order, more complex temporal organizations (see, for example, Deutsch & Feroe, 1981; Fraisse, 1978, 1982; M. R. Jones, 1976; Povel & Essens, 1985). Correspondingly, auditory graph research has suggested that X axis rhythmic context (in the form of clicks or beats) was more effective when the accents were placed at a density that was less than the density of the actual discrete data points (Smith & Walker, 2002, 2005). In other words, a click track helped little with temporal organization when the clicks coincided with each discrete data point (and the clicks offered only redundant information), but the click track was beneficial when it played, for example, once every two data points.

The manipulations of the current studies were such that both Y axis and X axis auditory context could not be held constant across the manipulations, as the contextual manipulations would potentially offer more benefit for some conditions than others. Instead, context was provided via the instructions and pre-experimental practice. Participants were familiarized with the scaling of data (i.e., which frequencies represented the highest and lowest prices of the day) as well as the mapping of stock price to the increasing and decreasing frequencies of tones. For all trials, participants were told the opening price of the stock (e.g., $50) as an additional contextual cue. In other words, although concurrent auditory context was not used, participants were given other contextual cues (that *could* be held constant across all conditions) to help them perform the graphing tasks.

2.1.5 Procedure and Task

Informed consent was obtained from participants before any procedures were performed. Due to the novelty of the display, all participants experienced a brief (approximately 10-15 min), self-paced, conceptual background presentation (see Smith & Walker, 2005; Walker & Nees, 2005b) that gave an overview of auditory graphs (e.g., "What is an auditory graph?" etc.), as well as instructions for the point estimation task. Participants were also given part-task practice on important component steps of the point estimation task during the introductory presentation, followed by a full set (16 trials) of whole-task practice with feedback on the experimental point estimation task. During these practice trials, participants heard auditory graph stimuli that were equivalent to the stimuli presented during test trials (with regard to the data density and the number of trend reversals), and the practice trials sampled the entire range of stimulus combinations that were used in the test trials. At no time during practice, however, did a participant hear any stimulus that was later used in a test trial. This was possible due to the counterbalancing of graphs with increasing and decreasing initial deflections across stimulus combinations. A person who would later experience linear increasing *experimental* trials (for a given combination of data density and trend reversals) was given linear decreasing examples (of the same given combination of density and trend reversals) during *practice* trials. Both part-task and whole-task training have been shown to improve performance on the point estimation task (Smith & Walker, 2005; Walker & Nees, 2005b), so the combined effects of both part-task and whole-task practice before testing helped to reduce effects of unfamiliarity with the display that might otherwise have been present in early test trials.

The 4 (density) x 4 (trend reversals) design resulted in 16 stimulus combinations. Participants therefore experienced 11 sets of 16 experimental trials, with each set consisting of randomly interleaved trials, one from each of the 16 data density and trend reversal stimulus combinations. Over the course of all 11 sets, participants were asked to identify the price of the stock for each hour of the trading day (8 a.m. – 6 p.m.) for each stimulus condition (16 combinations of data densities and trend reversals) in a random order.

Individual trials began with a visual text presentation of the test question (e.g., "What is the price of the stock at 10 a.m.?") followed by the presentation of the auditory graph. To provide a baseline reference, participants were told the opening price of the stock for each auditory graph. Participants were permitted to listen to the auditory graph as many times as needed before responding, and the next trial began after a response was recorded with the computer keyboard. Participants were given mandatory 5 min breaks between Sets 4 and 5 and between Sets 8 and 9. In an effort to promote engagement with the task, participants were given feedback about their performance for the set (as a whole) at the end of each of the 11 sets of experimental trials, but no *specific* feedback about individual experimental trial performance was given.

2.1.6 Dependent Variables

The primary dependent variable of Experiment 1 was operationally defined as the root mean squared (RMS) error in dollars of participants' responses to the point estimation trials for each of the 16 stimulus combinations. The mean number of times participants listened to graphs for each combination was a second dependent variable of interest.

2.2 Results

One outlier datum was removed from the analyses. One participant (presumably accidentally) responded with an estimated stock price of 878 dollars for the 10 a.m. question for the 3 trend reversals, 8 data points per second stimulus combination. Participants were explicitly told the data ranged up to only 106 dollars, and the participant gave no extreme responses for any other questions. This datum was not included in analyses, and the participant's RMS error for that stimulus combination was computed as a mean out of 10 trials instead of 11.

To check the assumption that the initial deflection of the graphed data (either increasing or decreasing, see Figures 1 and 2) had no impact on performance, a series of one-way analyses of variance (ANOVAs) was performed (for each combination of data densities and trend reversals) with the initial deflection of the graph as the independent variable. The Bonferroni procedure was used to protect family-wise alpha across this set of analyses. As was hypothesized, results for both dependent variables (RMS error and the mean number of times a graph was played) showed no significant difference with regard to the initial deflection of the graph for any stimulus combination (see Tables 1 and 2), and for further analyses the data were collapsed across the initial direction of the graph.

Table 1. *Experiment 1 (point estimation) ANOVA results for RMS error by condition, with direction of initial deflection (increasing or decreasing) as the independent variable.*

Condition				
Trend reversals	Points per second	*F* (1,30)	*p*	*M RMSE* difference (increasing-decreasing)
0	1	3.18	.09	-5.23
	2	2.41	.13	-4.84
	4	0.47	.48	-1.97
	8	0.19	.67	-1.66
1	1	0.67	.42	-2.60
	2	0.02	.88	-0.49
	4	0.01	.98	-0.11
	8	0.01	.92	0.33
2	1	0.01	.93	0.24
	2	2.56	.12	4.30
	4	0.42	.52	-1.83
	8	0.60	.44	-2.57
3	1	0.03	.86	0.53
	2	0.31	.58	-1.75
	4	0.05	.82	-0.65
	8	0.09	.77	-1.07

Note: No significant differences were found (Bonferroni family-wise corrected critical ∝ = .003), which supported the decision to collapse across the deflection variable for subsequent analyses.

Table 2. *Experiment 1 (point estimation) ANOVA results for mean number of times listened by condition with direction of initial deflection (increasing or decreasing) as the independent variable.*

Condition				
Trend reversals	Points per second	F (1,30)	p	M times listened difference (increasing-decreasing)
0	1	0.75	.39	-0.06
	2	2.87	.10	-0.10
	4	1.09	.30	0.07
	8	0.08	.78	-0.02
1	1	1.42	.24	-0.10
	2	0.90	.35	0.08
	4	0.20	.66	-0.03
	8	3.58	.07	0.12
2	1	0.26	.62	-0.06
	2	0.21	.65	0.06
	4	0.30	.59	-0.06
	8	0.23	.63	0.05
3	1	2.25	.14	0.21
	2	0.38	.54	-0.13
	4	0.02	.89	-0.01
	8	3.03	.09	-0.18

Note: No significant differences were found (Bonferonni family-wise corrected critical ∝ = .003), which supported the decision to collapse across the deflection variable for subsequent analyses.

Overall results for the collapsed data in the point estimation study are depicted in Figures 3 (RMS error) and 4 (mean number of times listened). Participants' RMS error scores and the mean number of times each type of graph was played were analyzed with a within-subjects multivariate analysis of variance (MANOVA). The MANOVA revealed significant main effects of both data density [*F* (6,26) = 4.68, $p = .002$, partial $\eta2 = .52$] and trend reversals [*F* (6,26) = 14.12, $p < .001$, partial $\eta2 = .77$]. The interaction of density with trend reversals was not significant [*F* (18,14) = 1.80, $p = .135$]. The results of the corresponding follow-up univariate tests are reported below; sphericity held for the RMS error dependent variable for both density (Mauchly's $W = .71$, $p = .07$) and trend reversal (Mauchly's $W = .69$, $p = .05$) effects. Sphericity assumptions were violated for the mean number of times listened for both density (Mauchly's $W = .38$, $p < .001$) and trend reversal (Mauchly's $W = .16$, $p < .001$) effects, thus a Huynh-Feldt correction was employed to guard against inflated alpha.

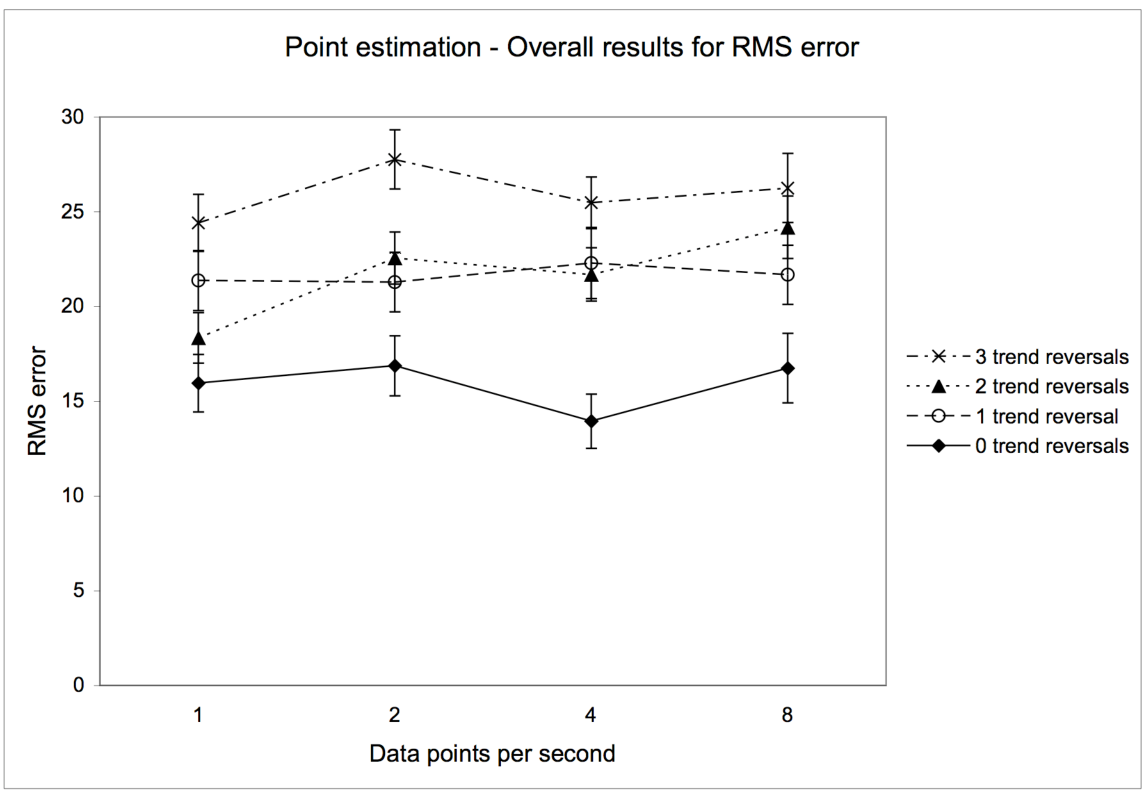

Figure 3. Overall results of the point estimation task (Experiment 1) for the *RMS* error dependent variable. Note that higher error indicates worse performance; error bars represent standard error.

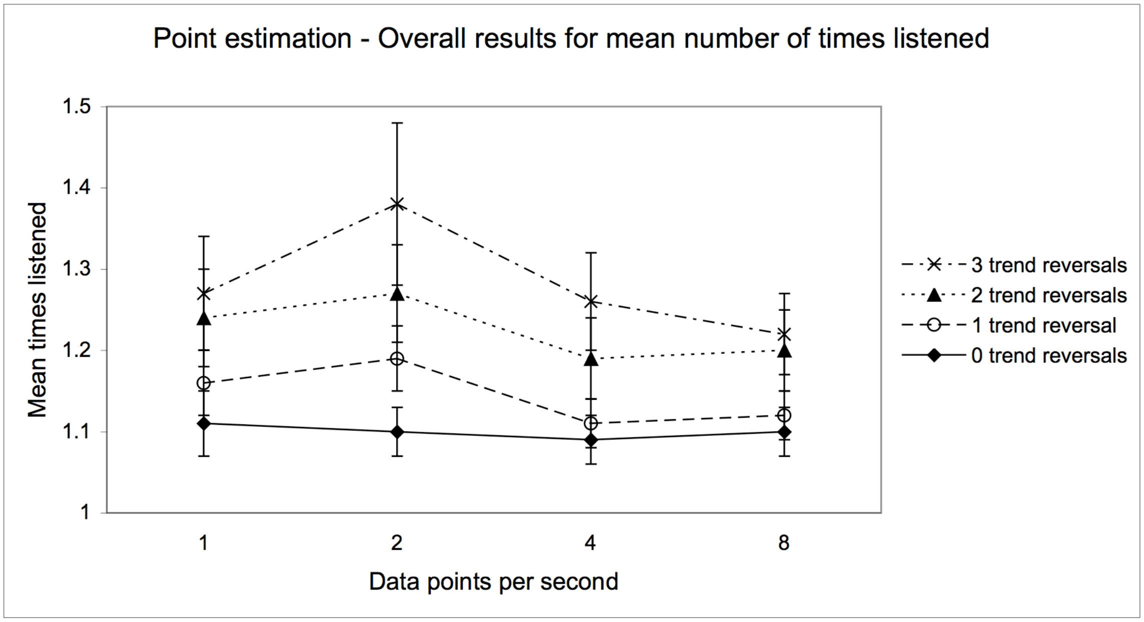

Figure 4. Overall results of the point estimation task (Experiment 1) for the mean number of times listened dependent variable. Error bars represent standard error.

Univariate follow-up tests for the main effect of density agreed with the results of the multivariate tests. The effect of density was significant on both *RMS* error scores [F (3,93) = 3.08, p = .03, partial η^2 = .09] and the mean number of times listened to each graph [F (2.01,62.17) = 6.86, p = .002, partial η^2 = .18]. Planned polynomial contrasts revealed a significant cubic component for *RMS* error [F (1,31) = 8.64, p = .006, partial η^2 = .22] and significant linear [F (1,31) = 5.37, p = .027, partial η^2 = .15] and cubic [F (1,31) = 15.13, p < .001, partial η^2 = .33] components for the mean number of times listened. Results for the main effect of data density (collapsed across trend reversals) are depicted in Figure 5.

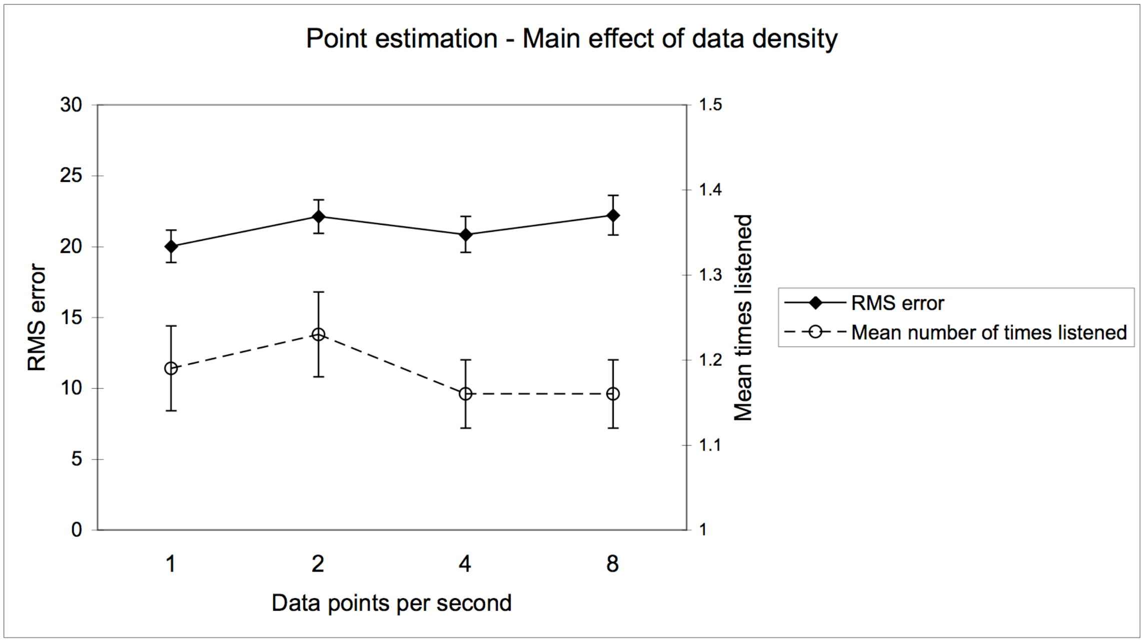

Figure 5. Results of the point estimation task (Experiment 1) for the significant main effect of data density on both dependent variables. Error bars represent standard error.

Similarly, univariate follow-up tests showed a significant main effect of trend reversals on both *RMS* error scores [F (3,93) = 41.09, p < .001, partial η^2 = .57] and the mean number of times listened to each graph [F (1.60,49.63) = 10.27, p < .001, partial η^2 = .25]. Planned polynomial contrasts on RMS error revealed significant linear [F (1,31) = 72.14, p < .001, partial η^2 = .70] and cubic [F (1,31) = 17.22, p < .001, partial η^2 = .36] components, while contrasts on the mean number of times listened showed a significant linear component [F (1,31) = 13.00, p = .001, partial η^2 = .30]. Results for the main effect of trend reversals (collapsed across data density) are depicted in Figure 6.

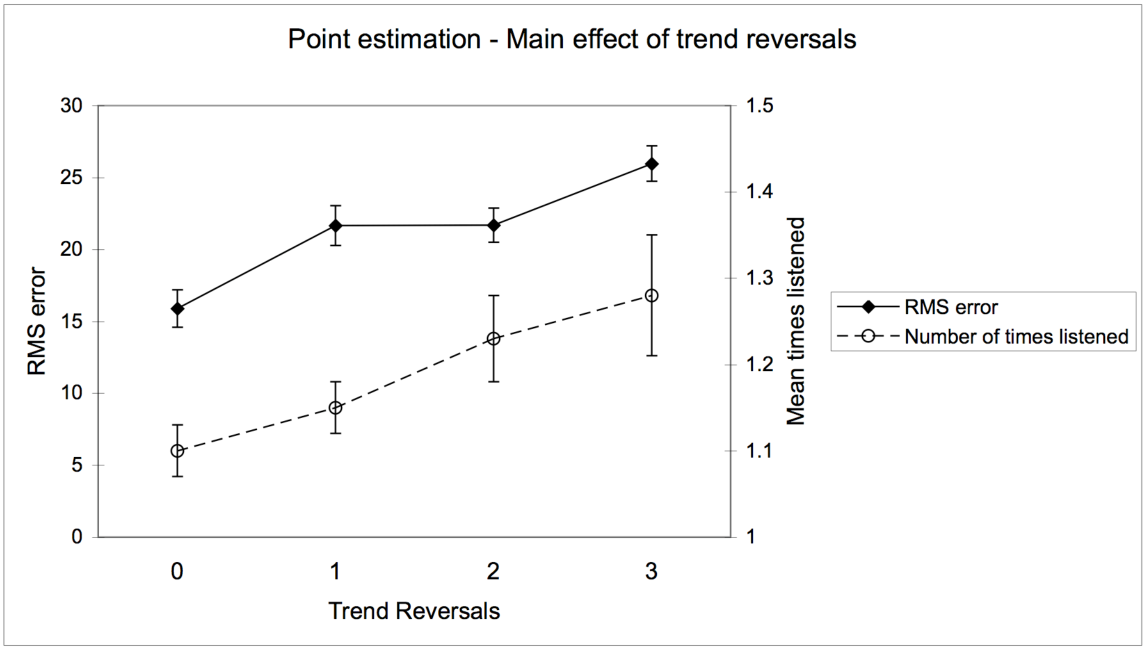

Figure 6. Results of the point estimation task (Experiment 1) for the significant main effect of trend reversals on both dependent variables. Error bars represent standard error.

An exploratory analysis was conducted to determine if performance changed across presentations of sets in time (collapsed across data density and trend reversals). As described above, each presentation set consisted of 16 randomly interleaved trials, each with one presentation of each stimulus combination for the independent variables. To investigate practice effects, another repeated measures MANOVA was employed with the *RMS* error for each set and the mean number of times listened to the graphs in the set as the dependent variables. The whole task practice set and the 11 sets of experimental trials were entered as the independent variable (which was, essentially, time). The MANOVA did not reveal a significant effect of set [F (22,10) = 1.71, p = .192]. For the univariate follow-ups, the sphericity assumption held for *RMS* error (Mauchly's W = .142, p = .869), but the mean number of times listened variable again required a Huynh-Feldt correction for violating sphericity assumptions (Mauchly's W = .00, $p < .001$). The univariate follow-up tests were not consistent with omnibus MANOVA and indicated effects of set on both *RMS* error [F (11,341) = 2.65, p = .003, partial η^2 = .079] and the mean number of times listened to each graph [F (4.29, 133.10) = 11.58, p < .001, partial η^2 = .27]. Follow-up t tests compared performance on the first set of experimental trials (Set 1) to performance on the last set of experimental trials (Set 11). For RMS error, performance during Set 1 was not significantly different than performance at Set 11 [t (1,31) = .72, p = .48]. The mean number of times listened, however, was significantly different for Set 1 as compared to Set 11 [t (1, 31) = 4.51, p < .001], with participants listening to each graph fewer times on average at the end of the experimental trials. Results across sets (i.e., time) in Experiment 1 are depicted in Figure 7.

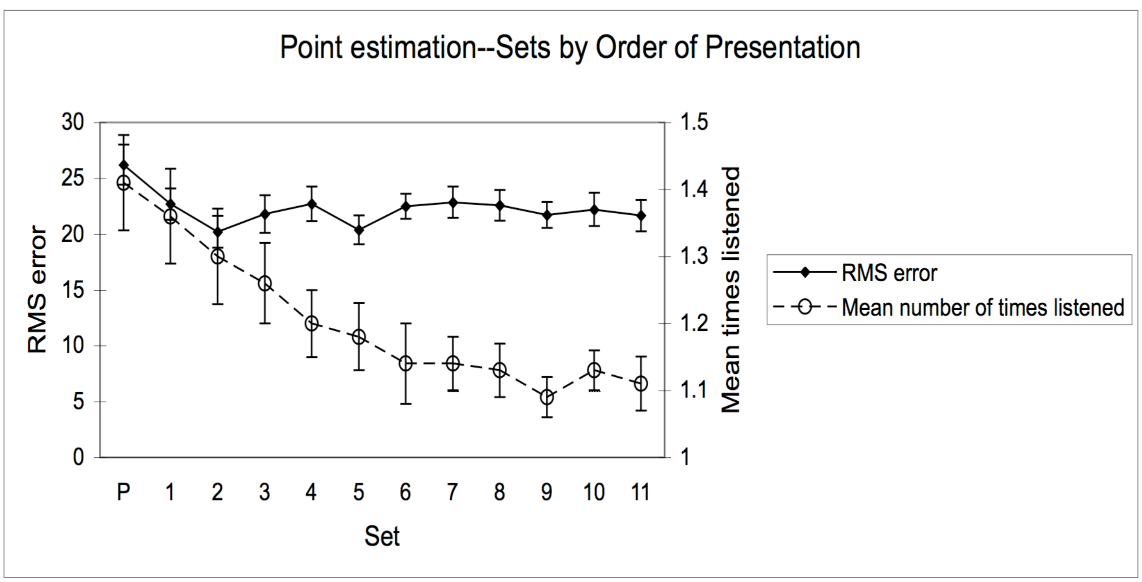

Figure 7. Results over time for the point estimation task (Experiment 1). The first set featured practice trials with feedback, while the rest of the sets were experimental trials with no specific feedback. Error bars represent standard error.

Finally, demographic variables related to musical experience and experience with stock trading, math, and business or economics courses were correlated with each other as well as with a gross measure of overall performance—the average RMS error across the entire study. Spearman's rho was employed due to the non-normal distribution of the demographic variables, and results are depicted in Table 3. Answers to some of the demographic questions correlated highly with other related demographic questions (as would be expected), but none of the demographic questions was significantly associated with performance on the point estimation task.

Table 3. *Experiment 1 (point estimation) Spearman correlation coefficients for demographic variables and average performance across all blocks.*

	RMS[1]	M1[2]	M2[3]	M3[4]	M4[5]	S1[6]	S2[7]	S3[8]
RMS	--							
M1	-.21	--						
M2	-.06	.43*	--					
M3	-.16	.92**	.37*	--				
M4	-.09	.90**	.34	.95**	--			
S1	.18	-.19	-.14	-.21	-.16	--		
S2	-.24	.17	-.18	.14	.16	.38*	--	
S3	.08	-.05	.04	.06	.02	.22	.39*	--

**p* < .05; ***p* < .01

[1]RMS: Average *RMS* error across all blocks in the study
[2]M1: Years having played a musical instrument
[3]M2: Age at which began playing a musical instrument
[4]M3: Years of formal musical training
[5]M4: Years experience reading musical notation
[6]S1: Years experience trading stocks or closely following the stock market
[7]S2: Number of college or advanced placement level math courses taken
[8]S3: Number of college or advanced placement level business/economics courses taken

2.3 Discussion

Regarding the main effect of data density for both dependent variables (Figure 5), a significant cubic component was found for both dependent measures such that, as the number of data points presented per second increased, performance first decreased (i.e., showed more error) at 2 points per second, then improved slightly (at 4 points per second), then faltered slightly again (at 8 points per second). Hypothesis 1a regarding data density predicted an increase in RMS error as the number of data points per second increased, and this hypothesis was not (or at least not fully) confirmed. Although the cubic component is difficult to account for as a psychological phenomenon, the mean differences constituting the cubic trend are small (-2.1 dollars, 1.27 dollars, and -1.63 dollars, respectively, as density changes from 1 to 2 points per second, 2 to 4 points per second, and 4 to 8 points per second). Furthermore, the main effect of density, while significant, accounted for only about 9% of the partialled variance in RMS error scores and 18% of the partialled variance in the mean number of times listened.

Results for the main effect of trend reversals for the point estimation task generally confirmed Hypothesis 1b that performance would decrease as the number of trend reversals increased. As depicted in Figure 6, the best performance (i.e., lowest RMS error and lowest mean number of times listening to the graphs) occurred when there were no trend reversals, and the worst performance occurred with 3 trend reversals in the data. The main effect of trend accounted for about 57% of the partialled variance in RMS error scores and about 27% of the partialled variance in the mean number of times listened.

Together these results suggest that the number of trend reversals played a large role in participants' ability to perform the point estimation task, while the number of data points presented per second had a significant but less substantial influence on performance outcomes. Performance (as measured by both RMS error and the mean number of times listened to each graph) clearly decreased as the number of trend reversals increased from 0 to 3. The data for RMS error showed a plateau for performance with 1 and 2 trend reversals (see Figure 6), which was reflected in the significant cubic contrast. Although this plateau is not evident in the data for the mean number of times listened, it warrants further investigation in future studies.

The results regarding data density may be of less practical than statistical significance, as the mean differences across all conditions were quite small and account for relatively little performance variance as compared to the trend reversals manipulation. Further research is needed to clarify the extent to which data density will impact performance with point estimation tasks in auditory graphs. The cubic performance trends found in the current study hint that there may be one or more rhythmic “sweet spots” that facilitate performance, but such a finding has not been suggested in past research and theory. Also, the small effect sizes and mean differences suggest that the main effect of data density on the point estimation task should be interpreted with caution.

The exploratory analyses to look for effects of set over time on the dependent variables produced conflicting results for the MANOVA, which indicated no effect of set, as compared to the follow-up univariate ANOVAS, which suggested effects of set on both dependent variables. Figure 7 plots these variables. A follow-up *t* test found no statistically significant differences over time for RMS error. This finding is consistent with past research, where Walker and Nees (2005b) showed that repeated exposure to the point estimation task in the absence of specific feedback for each trial did not result in improved performance as measured by RMS error. The follow-up *t* test for the mean number of times listened suggested that participants listened to the auditory graph stimuli fewer times on average at the end of the study. Although the current data can not conclusively explain this effect, the combination of significantly fewer mean times listened in Set 11 without a corresponding drop in accuracy (RMS error) in later trials suggests that participants did not need to hear the auditory graph stimuli as many times to maintain task performance by the end of the study.

CHAPTER 3

EXPERIMENT 2: TREND IDENTIFICATION

Experiment 2 examined the effects of data density and the number of data trend reversals on performance of a *trend identification task* with auditory graphs. A new sample of participants was recruited, but the apparatus, stimuli, and experimental manipulations were the same as those in Experiment 1. The only substantial difference in methodology between Experiments 1 and 2 was the task; participants in Experiment 2 were asked to identify local trends, and also the global shape or trend for each stimulus combination as described below.

3.1 Method

3.1.1 Participants

Participants (N = 32; 22 males and 10 females) were recruited from undergraduate psychology courses at the Georgia Institute of Technology, and none had participated in Experiment 1. Participants' mean age was 18.99 years (SD = 1.43), and the ages sampled ranged from 18 to 24 years old. Participants reported having played a musical instrument for an average of 4.78 (SD = 3.51) years, with 3 participants having never played an instrument and 16 having played an instrument for 5 years or more. Participants reported a mean of 3.94 (SD = 2.97) years of formal musical training (i.e., private or class instruction), and a mean of 4.91 (SD = 3.79) years of experience with reading musical notation. Participants had a mean of 0.47 (SD = 1.11) years of experience with stock trading or closely following the stock market. They also reported having taken a mean of 2.56 (SD = 1.76) college or advanced placement level math courses and an average of .97 (*SD* = 0.97) college or advanced placement level business or economics courses.

3.1.2 Procedure and Task: Differences between Experiments 1 and 2

The procedure and task for Experiment 2 were the same as Experiment 1 with a few notable exceptions. The introductory instructional materials and practice trials in Experiment 2 were exactly like those in Experiment 1, except the session was tailored to help participants to perform trend identification rather than point estimation.

Experiment 2 featured 11 sets of trials, and participants were asked to identify the local trend of the stock between all 10 successive hours of the trading day (e.g., “Did the price of the stock increase, decrease, or stay the same between 10 a.m. and 11 a.m.?”) for each stimulus condition in a random order. Participants were limited to three response choices: trend increased, trend decreased, or trend stayed the same.

The trend identification study had one exploratory trial per stimulus combination where, for each of the 16 combinations of data density and trend reversals, participants were asked to match the global shape (i.e., the overall trend—linear increasing, parabola, sine curve, etc.) of the data to its visual representation. Responses were limited to those 32 global shapes (of each combination of data density, trend reversals, and increasing or decreasing initial direction) that were in the entire stimulus set. Participants in Experiment 2 experienced 11 sets of 16 trials, with each set consisting of randomly interleaved trials.

Individual trials began with a visual text presentation of the test question (e.g., “Is the price of the stock increasing, decreasing, or staying the same between 9 a.m. and 10 a.m.?”) followed by the presentation of the auditory graph. To provide a baseline reference, participants were told the opening price of the stock for each auditory graph. Participants were permitted to listen to the auditory graph as many times as needed before responding, and the next trial began after a response had been recorded with the computer keyboard. Participants were given mandatory 5 min breaks between Sets 4 and 5 and between Sets 8 and 9. In an effort to promote engagement with the task, participants were given feedback about their performance for the set (as a whole) at the end of each of the 11 sets of trials, but no *specific* feedback about individual experimental trial performance was given.

3.1.3 Dependent Variables

The dependent variables of Experiment 2 were operationally defined as the percentage of correct responses on the local trend identification task for each of the 16 stimulus combinations. Again, the mean number of times participants listened to graphs for each set was a second dependent variable of interest. Finally, the data for the 16 exploratory global trend questions were transferred to a confusion matrix, and mistakes were examined for patterns of confusions.

3.2 Results

For Experiment 2, the first analyses again checked the assumption that the initial deflection of the graphed data (either increasing or decreasing, see Figures 1 and 2) had no impact on performance with a series of one-way ANOVAs for each combination of data densities and trend reversals. The initial deflection of the graph (either increasing or decreasing) was the independent variable. The Bonferroni procedure was again used to protect family-wise alpha across this set of analyses. As was hypothesized, results for both dependent variables (percent correct and the mean number of times a graph was played) showed no significant difference with regard to the initial deflection of the graph for any stimulus combination (see Tables 5 and 6), and for further analyses the data were collapsed across this variable.

Table 4. *Experiment 2 (trend identification) ANOVA results for mean percent correct by condition, with direction of initial deflection (increasing or decreasing) as the independent variable.*

Condition				
Trend reversals	Points per second	*F* (1,30)	*p*	*M* percent correct difference (increasing-decreasing)
0	1	0.01	.93	0.01%
	2	---	---	0.00
	4	2.14	.15	1.25
	8	0.12	.73	1.00
1	1	1.25	.27	4.70
	2	0.44	.52	- 3.10
	4	1.03	.32	- 3.70
	8	2.89	.09	- 8.00
2	1	0.31	.58	- 3.10
	2	5.78	.02	-10.00
	4	3.78	.06	- 8.10
	8	0.41	.53	- 3.40
3	1	4.38	.05	-12.50
	2	2.46	.13	- 8.75
	4	0.69	.41	5.00
	8	1.04	.32	6.20

Note: No significant differences were found (Bonferonni family-wise corrected critical ∝ = .003), which supported the decision to collapse across the deflection variable for subsequent analyses.

Table 5. *Experiment 2 (trend identification) ANOVA results for mean number of times listened by condition, with direction of initial deflection (increasing or decreasing) as the independent variable.*

Condition				
Trend reversals	Points per second	F (1,30)	p	M times listened difference (increasing-decreasing)
0	1	4.58	.04	-0.04
	2	1.90	.18	0.03
	4	3.43	.07	-0.05
	8	0.00	.97	-0.00
1	1	2.61	.12	-0.06
	2	0.05	.82	-0.01
	4	0.45	.51	-0.03
	8	0.73	.40	0.04
2	1	5.60	.03	0.14
	2	0.01	.92	-0.01
	4	0.05	.83	-0.02
	8	0.68	.42	-0.06
3	1	0.54	.47	0.05
	2	0.01	.91	-0.01
	4	0.01	.94	0.01
	8	0.70	.41	-0.06

Note: No significant differences were found (Bonferroni family-wise corrected critical $\propto$ = .003), which supported the decision to collapse across the deflection variable for subsequent analyses.

Participants' percent correct scores and the mean number of times each type of graph was played were analyzed with a within-subjects MANOVA, which revealed a significant main effect for the number of trend reversals [F (6,26) = 28.85, p < .001, partial η^2 = .87] but not data density [F (6,26) = 1.48, p = .223]. The interaction of density with trend reversals was not significant [F (18,14) = 2.14, p = .08]. Overall results for the trend identification study are depicted in Figures 8 (for percent correct) and 9 (for mean number of times listened).

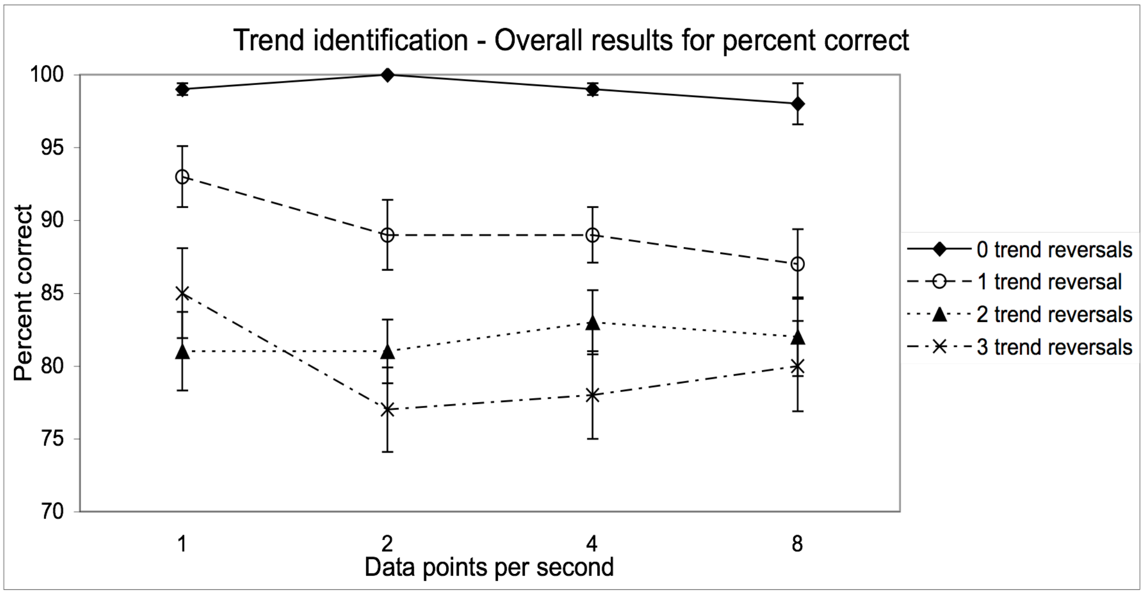

Figure 8. Overall results of the trend identification task (Experiment 2) for the percent correct variable. Error bars represent standard error.

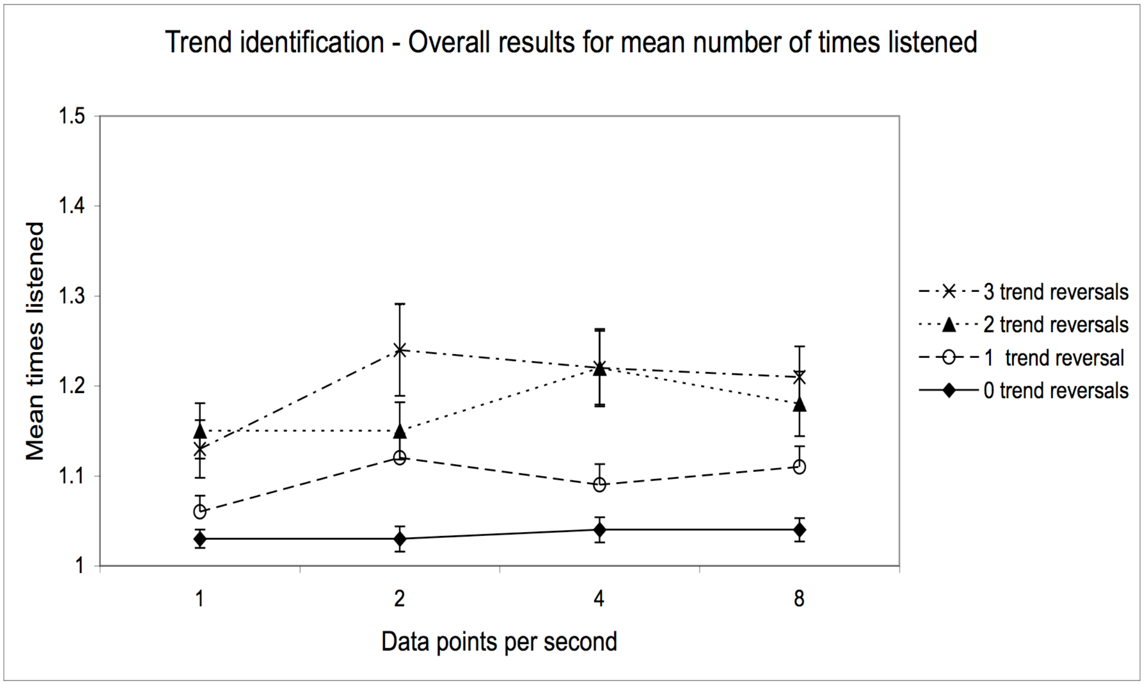

Figure 9. Overall results of the trend identification task (Experiment 2) for the mean number of times listened dependent variable. Error bars represent standard error.

The results of the corresponding univariate follow-up tests are reported below; sphericity assumptions held for only the percent correct dependent variable testing for the main effect of density (Mauchly's $W = .86$, $p = .47$). Sphericity assumptions were violated for the mean number of times listened dependent variable for the main effects of density (Mauchly's $W = .62$, $p = .01$), and trend reversals (Mauchly's $W = .47$, $p < .001$), and the density by trend reversals interaction (Mauchly's $W = .02$, $p < .001$). Sphericity assumptions were violated for the percent correct dependent variable for both the main effect of trend reversals (Mauchly's $W = .38$, $p < .001$) and for the density by trend reversals interaction (Mauchly's $W = .04$, $p < .001$). In all instances where sphericity was violated, a Huynh-Feldt correction was employed to guard against inflated alpha.

Univariate follow-up tests showed a significant main effect of density on the mean number of times participants listened to each type of graph [F (2.51,77.8) = 3.80, p = .02, partial η^2 = .11] but not on the percent correct [F (3,93) = 2.04, p = .11], which explains the nonsignificant omnibus MANOVA result for data density. Planned polynomial contrasts revealed a significant linear component for the mean number of times listened across densities [F (1,31) = 5.21, p = .03, partial η^2 = .14]. Results for the main effect of data density on the mean number of times listened (collapsed across trend reversals) are depicted in Figure 10.

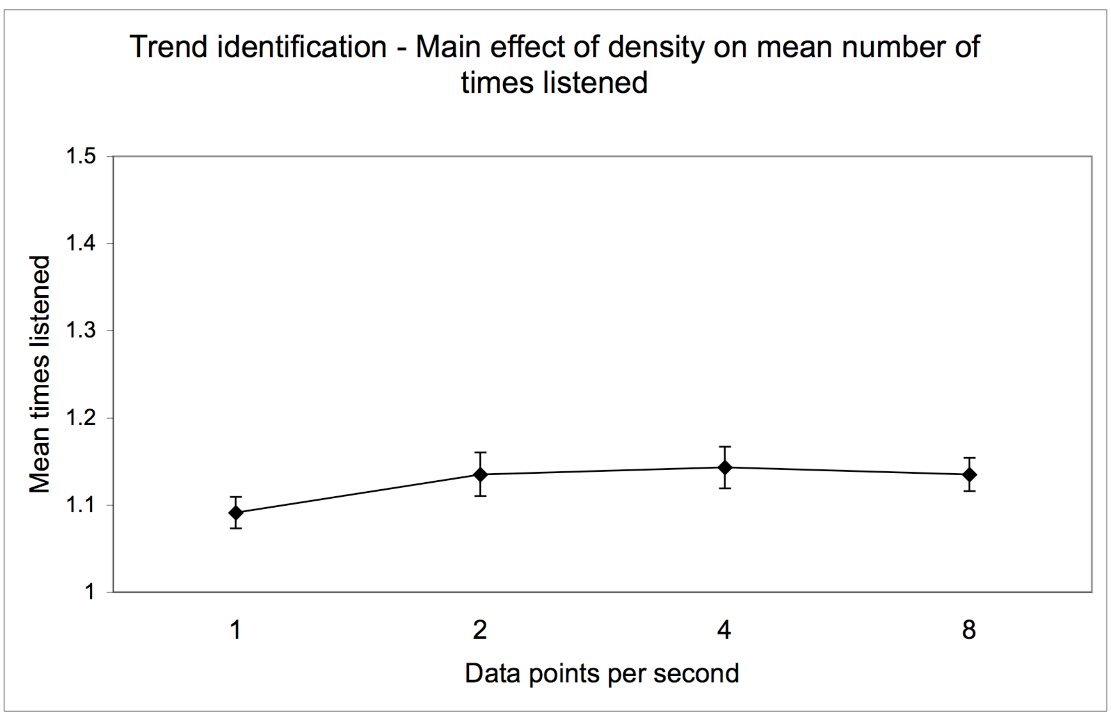

Figure 10. Results of the trend identification task (Experiment 2) for the significant main effect of data density on the mean number of time listened to each graph. Error bars represent standard error.

Univariate follow-up tests for the main effect of trend reversals showed a significant effect on both percent correct scores [F (1.93,58.81) = 60.78, $p < .001$, partial $\eta^2 = .66$] and the mean number of times listened to each graph [F (2.13,65.97) = 23.57, $p < .001$, partial $\eta^2 = .43$]. Planned polynomial contrasts on the percent correct revealed significant linear [F (1,31) = 82.86, $p < .001$, partial $\eta^2 = .73$] and quadratic [F (1,31) = 22.57, $p < .001$, partial $\eta^2 = .42$] components, while contrasts on the mean number of times listened showed a significant linear component [F (1,31) = 34.78, $p < .001$, partial $\eta^2 = .53$]. Results for the main effect of trend reversals (collapsed across data density) on both dependent measures are depicted in Figure 11.

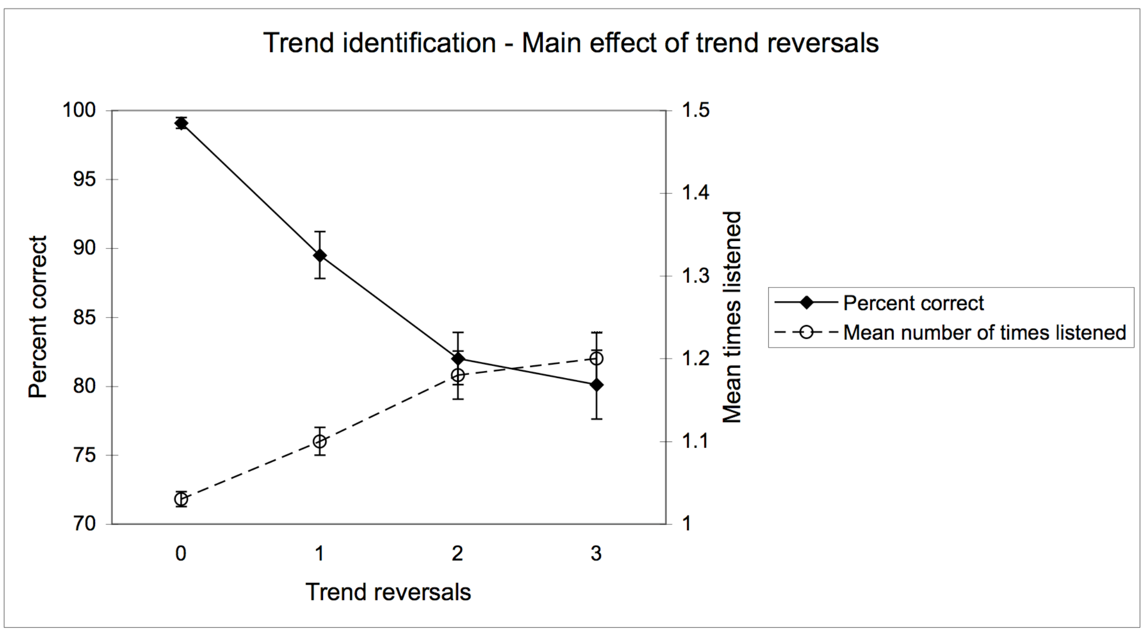

Figure 11. Results of the trend identification task (Experiment 2) for the significant main effect of trend reversals on both dependent variables. Error bars represent standard error.

Univariate follow-up tests also revealed a significant interaction of the data density and trend reversals manipulations for percent correct [F (7.06,218.7) = 2.11, p = .04, partial η^2 = .06], but not for the mean number of times listened to each graph stimulus [F (6.46,200.19) = 1.478, p = .18], hence the nonsignificant MANOVA interaction term. Again, the interaction for percent correct is depicted in Figure 8. Repeated measures interaction contrasts revealed that the interaction effect resulted in different patterns for the trend reversals variable for 1 data point per second as compared to 2 data points per second for the density variable. Interaction contrasts revealed significant interactions for 0 trend reversals as compared to 1 trend reversal as density increased from 1 to 2 data points per second [F (1,31) = 6.90, p = .013, partial η^2 = .18], and also for 2 trend reversals as compared to 3 trend reversals as density increased from 1 to 2 data points per second [F (1,31) = 7.95, p = .008, partial η^2 = .20]. The interactions are observed in the upper left and lower left corners of Figure 8.

Like in Experiment 1, a final, exploratory set of analyses was conducted to determine if performance changed across presentations of sets in time (collapsed across data density and trend reversals). As described above, each presentation set consisted of randomly interleaved trials, each with one presentation of each stimulus combination for the independent variables. The MANOVA examined the percent correct for each set and the mean number of times listened to the graphs in the set as dependent variables, and the global trend matching questions were not included in this analysis. The practice set and the 11 sets of experimental trials were entered as the independent variable. The MANOVA did not reveal a significant effect of set [F (22,10) = 2.26, p = .09]. For the univariate follow-up tests, the sphericity assumption was violated for both the percent correct (Mauchly's $W = .03$, $p = .02$) and the mean number of times listened (Mauchly's $W = .03$, $p = .02$), and a Huynh-Feldt correction was again required. The univariate follow-up ANOVAs indicated effects of set (i.e., time) on the mean number of times listened [F (2.68,83.01) = 8.15, $p < .001$, partial $\eta^2 = .21$] but not on the percent correct [F (10.11,313.19) = 1.78, $p = .06$]. A follow-up t test compared performance on the first set of experimental trials (Set 1) to performance on the last set of experimental trials (Set 11) for the mean number of times listened. Performance was significantly different for Set 1 as compared to Set 11 [t (1, 31) = 4.03, $p < .001$], with participants listening to each graph fewer times on average at the end of the experimental trials. Results across sets (i.e., time) in Experiment 2 are depicted in Figure 12.

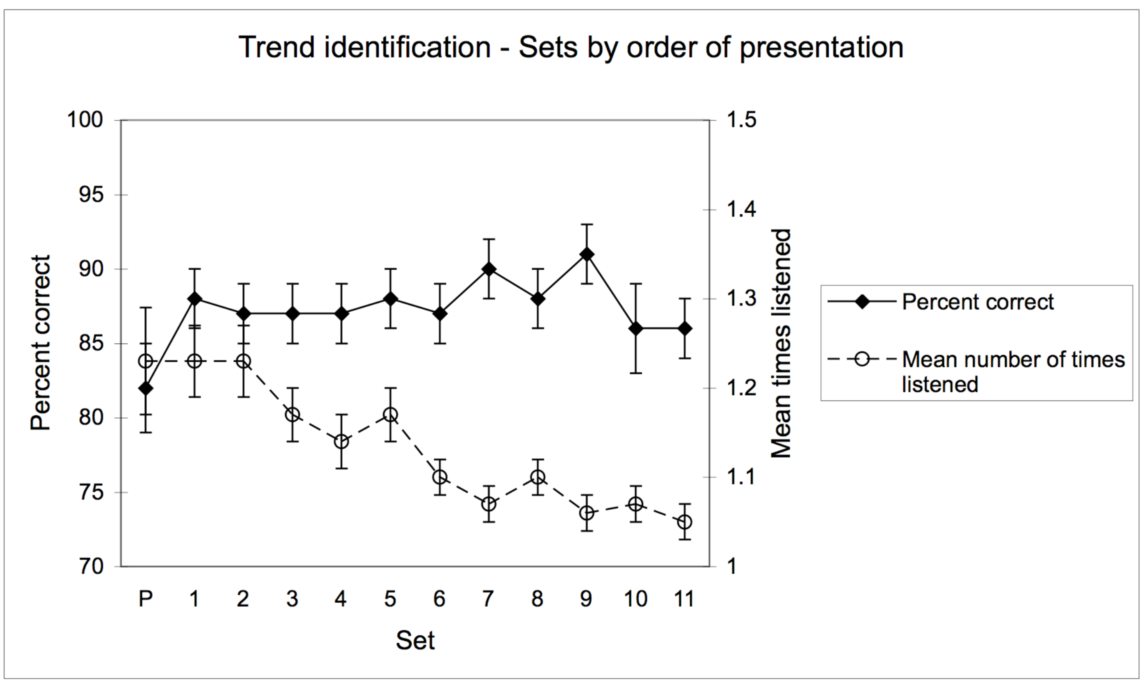

Figure 12. Results over time for the trend identification task (Experiment 2). The first set featured practice trials with feedback, while the rest of the sets were experimental trials with no specific feedback. Error bars represent standard error.

Demographic variables related to musical experience and experience with stock trading, math, and business or economics courses were correlated with each other as well as with a gross measure of overall performance—the average percent correct across the entire study. Spearman's rho was employed due to the non-normal distribution of the demographic variables, and results are depicted in Table 6. Again, answers to some of the demographic questions correlated highly with other related demographic questions (as would be expected), but none of the demographic questions was significantly associated with performance on the trend identification task.

Table 6. *Experiment 2 (trend identification) Spearman correlation coefficients for demographic variables and average performance across all blocks.*

	PER[1]	M1[2]	M2[3]	M3[4]	M4[5]	S1[6]	S2[7]	S3[8]
PER	--							
M1	.13	--						
M2	-.05	-.18	--					
M3	.12	.87**	-.09	--				
M4	.11	.88**	-.15	.82**	--			
S1	-.25	.17	-.01	.05	-.03	--		
S2	.03	.15	-.10	.19	.07	.37*	--	
S3	-.13	-.02	.15	-.02	-.05	.20	.42*	--

*$p < .05$; **$p < .01$
[1]PER: Average percent correct across all blocks in the study
[2]M1: Years having played a musical instrument
[3]M2: Age at which began playing a musical instrument
[4]M3: Years of formal musical training
[5]M4: Years experience reading musical notation
[6]S1: Years experience trading stocks or closely following the stock market
[7]S2: Number of college or advanced placement level math courses taken
[8]S3: Number of college or advanced placement level business/economics courses taken

The confusion matrix for the global trend question (matching the stimulus sound to a visual depiction of the overall graph shape) for each type of graph is shown in Appendix A. The matrix suggested that overall performance was high for many stimulus combinations, and well above chance for all graphs. Errors, in general, occurred with a high frequency one cell above the diagonal—a pattern that, for this matrix, suggested that participants tended to underestimate the density of sounds (and thus the overall number of data points) when choosing a visual graph to match the auditory graph that was heard. Also, moving from left to right in the matrix (i.e., from 0 to 3 trend reversals), the confusions tended to spread further away from the diagonal. This general observation suggested that, as trend reversals increased, participants were more likely to confuse the sound they heard with a visual depiction that had a different number of trend reversals altogether. These observations were consistent with past work (Bonebright et al., 2001) that found overall high accuracy for matching auditory graphs to their visual counterparts with a tendency for more amorphous data sets to pose more matching difficulty.

3.3 Discussion

Figure 10 depicts the significant main effect of data density on the mean number of times listened to each graph. Polynomial contrasts confirmed a significant linear trend component, which is largely reflected in the increase in the number of times participants listened to a graph as data density increased from 1 data point per second to 2 data points per second. The graph suggests that participants needed slightly fewer times listening to a graph that had one data point per second as compared to other density conditions. This finding regarding a main effect for density accounted for about 11% of the partialled variance in the mean number of times listened to each graph type, and the mean difference in the number of times listened for 1 point per second as compared to 2 points per second was only .044. (To put this in perspective, the minimum number of times a person could listen to each graph was once per graph across 16 trials in a set. A person who listened to one graph one extra time would have a mean number of times listened = 17/16 = 1.0625, a .0625 difference for having listened to one graph one time beyond the minimum in a set of 16 trials). No main effect of data density was found for the percent correct measure.

A main effect of the number of trend reversals was found for both dependent measures, and these are depicted in Figure 11. Both variables were found to have significant linear trend components, with the percent correct measure also featuring a significant quadratic component. Figure 11 shows generally parallel but opposite patterns. Performance on the task decreased concurrently for both measures, as reflected in simultaneous decreases in percent correct and increases in the mean number of times listened to the graphs as the number of trend reversals increases.

The finding of a main effect of trend reversals for the percent correct variable, however, should be interpreted in light of the significant interaction of data density and trend reversals for the percent correct. (No such interaction was present for the mean number of times listened). The interaction is depicted in Figure 8, and interaction contrasts revealed that the interaction was a function of differences in the pattern of the trend reversals variable from data density levels of 1 data point per second to 2 data points per second. While performance with 0 trend reversals remained near ceiling moving from 1 data point per second to 2 data points per second (and across all densities, for that matter), the interaction occurred such that the introduction of 1 trend reversal showed a decline in performance at 2 data points per second. Interestingly, the other significant interaction contrast suggested that, comparing 2 trend reversal performance to 3 trend reversal performance from 1 point per second to 2 points per second, performance for graphs with 3 trend reversals was better (than performance with only 2 trend reversals) at 1 point per second. Performance with 3 trend reversal graphs then dropped relatively sharply at 2 data points per second. This was contrary to hypothesis 2b, which predicted an interaction characterized by performance with 3 trend reversals and 1 data point per second as the worst stimulus combination for the trend identification task. The Nyquist (1928) theorem, therefore, did not hold for these data and this task.

The interaction of data density with trend reversals, coupled with the linear increase in the number of times participants tended to listen to a graph as data density increased from 1 data point per second to 2 data points per second, suggest that the change from 1 to 2 data points per second was problematic, particularly for conditions where graphs featured 1 or 3 trend reversals. For 0 trend reversals, performance remained at ceiling across remained at ceiling across manipulations of data density, while 2 trend reversals in the data showed little change as density increased from 1 to 2 data points per second. In some regards, the slowest rate of presentation seems to have compensated for the generally more difficult stimuli with 3 trend reversals, but it is unclear why the same pattern was not observed for data with 2 trend reversals.

Like in Experiment 1, the exploratory analysis to look for effects of set over time on the dependent variables produced conflicting results for the MANOVA, which indicated no effect of set, as compared to the follow-up univariate ANOVAS, which suggested effects of set on only the mean number of times listened. Follow-up *t* tests suggested that the percent correct for early trials was not significantly different for early trials (Set 1) as compared to later trials (Set 11), but participants listened to the auditory graph stimuli fewer times on average in the later trials. These results for the trend identification study parallel the findings from the same analyses for the point estimation study. Performance accuracy did not change over time, but participants needed fewer presentations of the auditory graph stimuli to maintain the same level of performance for later trials as compared to early trials.

CHAPTER 4

GENERAL DISCUSSION

This study's results regarding trend reversals suggested that, as was implied in theories by both M.R. Jones (1976) and Deutsch and Feroe (1981), an auditory graph with linear increasing or decreasing data is relatively easy to comprehend. When data are represented with sound, some features of data—particularly if those features are parsimonious with regard to trend—may be easily perceived and perhaps should be characterized as auditory versions of the emergent patterns discussed by Sanderson et al. (1989) in their discussion of visual displays. This finding similarly parallels the predictions made in Kosslyn's (1989) theory of visual graph comprehension, namely that simple patterns may be perceived automatically and efficiently in graphical representations.

The manipulation of trend reversals in the current study did not control for changes in the frequency intervals between queried data points as the number of trend reversals increased, as the scaling factor was instead held constant across stimuli. Although M.R. Jones' (1976) rhythmic theory directly predicts that complexity (i.e., perceivability) of an auditory sequence declines as larger and less constant intervals between pitches are introduced, other work (i.e., Dowling, 1978) has suggested that the processing of contour (i.e., trend changes) and intervals proceeds independently. A straightforward and imperative follow-up to the current study should disentangle the respective roles of trend reversals and frequency intervals between tones by adjusting data scaling to reflect equal and constant interval changes across different numbers of trend reversals.

Although more research is needed to isolate the exact data properties (intervals changes, trend changes, or both) whereby more trend reversals generally tended to result in worse auditory graph comprehension in the current studies, current practices (and current software, etc.) for making graphs from sound have never before considered that trend patterns in the data-to-be-represented may be a critical factor influencing graph comprehension. The current findings suggest that, holding time and frequency scaling factors constant, performance will generally be impacted negatively as the number of trend reversals increases. Performance was at or near the ceiling for the trend identification task across all data densities when the data had no trend reversals, and performance generally declined from ceiling levels as trend reversals were added. The same general pattern of findings was evident in the point estimation study as well. This does not mean that the introduction of trend reversals will render auditory graphs unusable altogether, and this finding parallels theoretical predictions regarding the comprehension of traditional visual graphs (e.g., Carswell et al., 1993; Carswell & Ramzy, 1997). It seems that the simplest auditory graphs are characterized by unidirectional, constant frequency change, which elicit Gestalt perceptual good form due to temporal proximity, frequency proximity, and expectations of good continuation of the frequency pattern. The same types of perceptual Gestalts (e.g., spatial proximity, temporal proximity, good continuation, etc.) likely contribute to the automatic perception of patterns in visual graphs as described by Kosslyn (1989), which have also been characterized as *emergent* features or patterns (Sanderson et al., 1989). It seems plausible that, as data relations become more complex (i.e., as they depart from simple linear increasing or decreasing relations, etc.), the comprehension of a graph, either visual or auditory, becomes more difficult. In extreme cases where data are random and characterized by drastic changes in Y axis values across small changes in X axis values, no auditory or visual pattern will emerge from graphical representations in any modality, because no pattern is present.

Current practices in auditory graphing generally start with choosing a scaling factor (e.g., 100 Hz -2000 Hz), then sonifying the data within that scaling factor. The dilemma regarding the choice of a scaling factor becomes more problematic when data are sonified in real time and a priori maximum and minimum data points are unknown. More research is needed to determine if, as predicted by rhythmic theory (M. R. Jones, 1976), a proportional scaling change to reduce frequency intervals could be used to alleviate auditory graph comprehensions problems for data with large, rapid changes in a short period of time. Auditory Y axis context in the form of reference tones (see Smith & Walker, 2002, 2005), which was not used in the current study, may also aid in the performance of auditory graphing tasks for data with more frequent trend changes. Furthermore, sonification researchers have yet to look at performance for highly trained listeners. Research has suggested that a brief training session helps naïve auditory graph listeners (e.g. Smith & Walker, 2005; Walker & Nees, 2005b), and more extensive investigations of training and practice with auditory graphs may show that training interventions also ameliorate some of the detriments in performance observed in the current study as trend reversals in data increased. The data of the current study, whose design and analyses were not tailored to examine either practice or training, showed no significant differences in accuracy for either study comparing early trials to later trials, but the analyses of performance over time in both studies suggested that participants needed to hear stimuli fewer times by the end of the study to maintain the same level of performance. Interestingly, the general patterns of the decreasing functions for mean times listened were more or less congruent with traditional practice functions (see Newell & Rosenbloom, 1981). Research regarding training with auditory graphs (e.g., Walker & Nees, 2005b) has suggested that considerable improvements over time might be possible if specific feedback for individual trials was introduced.

The relatively small impact of data density as compared to trend reversals in the current study does not necessarily suggest that the timing of data points will have only a small effect on auditory graph comprehension across all reasonable choices of data density per unit of time. The patterns in the current study were isochronous, with regular timing of data points occurring within a given graph. Clearly not all graphical data lend themselves to regular X axis spacing (e.g., scatterplots, etc.) The theories of Martin (1972) and M.R. Jones (1976) both predict that non-isochronous data densities, occurring as a result of data that are irregular on an X axis value or missing data points, etc., will be problematic for auditory pattern perception. Irregular timing of data, particularly data sets whose timing as auditory graphs follow no predictable hierarchical pattern, will not allow for attention to be focused based on expectancies. Smith and Walker (2002; Smith & Walker, 2005), however, have shown that X axis auditory context—in the form of regular beats or clicks—can help auditory graph users with the temporal organization of data. Such contextual cues may ameliorate the potential problems that could otherwise persist when sonifying non-isochronous data.

Auditory graphs have a great potential to improve data accessibility for blind students and scientists as well as sighted people. This study investigated the role of two important attributes of data--the density and the number of trend reversals. Results suggested that both data density and the number of trend reversals impacted the graph complexity and ease of data comprehension for auditory graph users across point estimation and trend identification tasks, with a generally greater effect attributable to the role of the number of trend reversals. The results of the current study can be used to help predict stimulus complexity for future studies of training for auditory graph users, as examinations of learning will require auditory graph stimuli that range in difficulty. Furthermore, the current study has suggested that theories of auditory pattern perception (e.g., M. R. Jones, 1976; Martin, 1972) may offer useful insights in understanding how the data in an auditory graph can be presented such that relations are simple and easy to perceive and understand.

APPENDIX A: CONFUSION MATRIX FOR GLOBAL TREND MATCHING QUESTION

Trend reversals:	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3
Points per second:	1	2	4	8	1	2	4	8	1	2	4	8	1	2	4	8	1	2	4	8	1	2	4	8	1	2	4	8	1	2	4	8
Graph	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
1	93	25											7																			
2	7	69	44																													
3			50	18																												
4			6	76																				6								
5					100	25																			13							
6						69	44	7																								
7						6	50	47																								
8							6	47																								
9		6							88	13											7	6							12	6		
10									6	69	31							6			7									6		
11				6						6	56	27																				
12											6	6																				
13									6	6			72	25				6			7	6										
14													7	69	32											13						
15															56	30																
16																70																
17																	90	19											12			
18															6			56	38	7			13							6		
19											6				6			6	56	12			6								13	
20																				60												7
21													7								72	38	6		27	6				6		
22										6												50	19			13	6					
23																							56	29		6	19					
24												13									7			59				10				
25																									53	6						
26																		6							7	50	13		6			
27																			6								62	20			6	7
28																				7				6				60				7
29													7	6			10												58	13	6	
30																										6			12	57	25	
31																				7										6	50	20
32																				7								10				59

REFERENCES

American Psychological Association. (2001). *Publication Manual of the American Psychological Association* (5th ed.). Washington, DC: American Psychological Association.

Bonebright, T. L., Nees, M. A., Connerley, T. T., & McCain, G. R. (2001). Testing the effectiveness of sonified graphs for education: A programmatic research project. *Proceedings of the International Conference on Auditory Display (ICAD2001)* (pp. 62-66), Espoo, Finland.

Bregman, A. S. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound.* Cambridge, MA: MIT Press.

Brewster, S. (1997). Using non-speech sound to overcome information overload. *Displays*, 17, 179-189.

Brewster, S. (2002). Overcoming the lack of screen space on mobile computers. *Personal and Ubiquitous Computing*, 6(3), 188-205.

Brewster, S., & Murray, R. (2000). Presenting dynamic information on mobile computers. *Personal Technologies*, 4(4), 209-212.

Brown, L. M., Brewster, S., & Riedel, B. (2002). Browsing modes for exploring sonified line graphs. *Proceedings of the 16th British HCI Conference*, London, UK.

Brown, L. M., & Brewster, S. A. (2003). Drawing by ear: Interpreting sonified line graphs. *Proceedings of the International Conference on Auditory Display (ICAD2003)* (pp. 152-156), Boston, MA.

Brown, L. M., Brewster, S. A., Ramloll, R., Burton, M., & Riedel, B. (2003). Design guidelines for audio presentation of graphs and tables. *Proceedings of the International Conference on Auditory Display (ICAD2003)* (pp. 284-287), Boston, MA.

Brown, M. L., Newsome, S. L., & Glinert, E. P. (1989). An experiment into the use of auditory cues to reduce visual workload. *Proceedings of the ACM CHI 89 Human Factors in Computing Systems Conference (CHI 89)* (pp. 339-346).

Butler, D. L. (1993). Graphics in psychology: Pictures, data, and especially concepts. *Behavior Research Methods, Instruments & Computers*, 25(2), 81-92.

Carpenter, P. A., & Shah, P. (1998). A model of the perceptual and conceptual processes in graph comprehension. *Journal of Experimental Psychology: Applied*, 4(2), 75-100.

Carswell, C. M. (1992). Reading graphs: Interactions of processing requirements and stimulus structure. In B. Burns (Ed.), *Percepts, Concepts and Categories: The Representation and Processing of Information.* (pp. 605-645). North-Holland: Elsevier.

Carswell, C. M., Emery, C., & Lonon, A. M. (1993). Stimulus complexity and information integration in the spontaneous interpretation of line graphs. *Applied Cognitive Psychology*, 7, 341-357.

Carswell, C. M., & Ramzy, C. (1997). Graphing small data sets: Should we bother? *Behaviour & Information Technology*, 16(2), 61-71.

Childs, E. (2005). Auditory graphs of real-time data. *Proceedings of the International Conference on Auditory Display (ICAD2005)*, Limerick, Ireland.

Deutsch, D., & Feroe, J. (1981). The internal representation of pitch sequences in tonal music. *Psychological Review*, 88(6), 503-522.

Dowling, W. J. (1978). Scale and contour: Two components of a theory of memory for melodies. *Psychological Review*, 85(4), 341-354.

Drake, C., & Botte, M. C. (1993). Tempo sensitivity in auditory sequences: Evidence for a multiple look model. *Perception & Psychophysics*, 54(3), 277-286.

Flowers, J. H., Buhman, D. C., & Turnage, K. D. (1997). Cross-modal equivalence of visual and auditory scatterplots for exploring bivariate data samples. *Human Factors*, 39(3), 341-351.

Flowers, J. H., Buhman, D. C., & Turnage, K. D. (2005). Data sonification from the desktop: Should sound be part of standard data analysis software? *ACM Transactions on Applied Perception*, 2(4), 467-472.

Flowers, J. H., & Hauer, T. A. (1992). The ear's versus the eye's potential to assess characteristics of numeric data: Are we too visuocentric? *Behavior Research Methods, Instruments & Computers*, 24(2), 258-264.

Flowers, J. H., & Hauer, T. A. (1993). "Sound" alternatives to visual graphics for exploratory data analysis. *Behavior Research Methods, Instruments & Computers*, 25(2), 242-249.

Flowers, J. H., & Hauer, T. A. (1995). Musical versus visual graphs: Cross-modal equivalence in perception of time series data. *Human Factors*, 37(3), 553-569.

Fraisse, P. (1978). Time and rhythm perception. In M. P. Friedman (Ed.), *Handbook of Perception. Volume VIII: Perceptual Coding* (pp. 203-254). New York: Academic Press.

Fraisse, P. (1982). Rhythm and tempo. In D. Deutsch (Ed.), *The Psychology of Music* (pp. 249-180). New York: Academic Press.

Friel, S. N., Curcio, F. R., & Bright, G. W. (2001). Making sense of graphs: Critical factors influencing comprehension and instructional applications [Electronic version]. *Journal for Research in Mathematics*, 32(2), 124-159.

Frysinger, S. P. (2005). A brief history of auditory data representation to the 1980s. *Proceedings of the International Conference on Auditory Display (ICAD2005)*, Limerick, Ireland.

Gillan, D. J., & Lewis, R. (1994). A componential model of human interaction with graphs: I. Linear regression modeling. *Human Factors*, 36(3), 419-440.

Gillan, D. J., Wickens, C. D., Hollands, J. G., & Carswell, C. M. (1998). Guidelines for presenting quantitative data in HFES publications. *Human Factors*, 40(1), 28-41.

Gobbo, C. (1994). On children's understanding of an economic concept: The role of graphics in evaluation. In W. Schnotz & R. W. Kulhavy (Eds.), *Comprehension of Graphics* (pp. 227-249). Amsterdam: Elsevier Science.

Grudin, J. (2001). Partitioning digital worlds: Focal and peripheral awareness in multiple monitor use. *Proceedings of the 2001 SIGCHI conference on Human Factors in Computing Systems* (pp. 458-465), Seattle, WA.

Jones, M. R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review*, 83(5), 323-355.

Jones, M. R., Johnston, H. M., & Puente, J. (2006). Effects of auditory pattern structure on anticipatory and reactive attending. *Cognitive Psychology*, 53, 59-96.

Jones, R. W., & Careras, I. E. (1996). The empirical investigation of factors affecting graphical visualizations. *Behavior Research Methods, Instruments & Computers*, 28(2), 265-269.

Kosslyn, S. M. (1989). Understanding charts and graphs. *Applied Cognitive Psychology*, 3(3), 185-225.

Kramer, G. (1994). An introduction to auditory display. In G. Kramer (Ed.), *Auditory Display: Sonification, Audification, and Auditory Interfaces* (pp. 1-78). Reading, MA: Addison Wesley.

Kubovy, M. (1981). Concurrent pitch segregation and the theory of indispensable attributes. In M. Kubovy & J. Pomerantz (Eds.), *Perceptual Organization* (pp. 55-99). Hillsdale, NJ: Lawrence Erlbaum.

Kubovy, M. (1988). Should we resist the seductiveness of the space:time::vision:audition analogy? *Journal of Experimental Psychology: Human Perception & Performance*, 14, 318-320.

Lee, J. C. K., & Gerber, R. (1999). Hong Kong students' perceptions of graphs, charts, and maps. *Scandinavian Journal of Educational Research*, 43(1), 19-40.

Levitin, D. J. (1999). Memory for musical attributes. In P. Cook (Ed.), *Music, Cognition, and Computerized Sound: An Introduction to Psychoacoustics.* (pp. 209-227). Cambridge, MA: MIT Press.

Liu, B., Salvendy, G., & Kuczek, T. (1999). The role of visualization in understanding graphics. *International Journal of Cognitive Ergonomics*, 3(4), 289-305.

Mankoff, J., Fait, H., & Tran, T. (2005). Is your web page accessible? A comparative study of methods for assessing web page accessibility for the blind. *Proceedings of the ACM CHI '05 Conference on Human Factors Computing Systems (CHI 05)*, Portland, OR.

Mansur, D. L., Blattner, M. M., & Joy, K. I. (1985). Sound graphs: A numerical data analysis method for the blind. *Journal of Medical Systems*, 9(3), 163-174.

Martin, J. G. (1972). Rhythmic (hierarchical) versus serial structure in speech and other behavior. *Psychological Review*, 79(6), 487-509.

Massachusetts Institute of Technology. (n.d.). MIT web accessibility guidelines. Retrieved April 10, 2006, from http://web.mit.edu/atic/www/accessibility/developweb.html

McAuley, J. D., & Jones, M. R. (2003). Modeling effects of rhythmic context on perceived duration: A comparison of interval and entrainment approaches to short-interval timing. *Journal of Experimental Psychology: Human Perception & Performance*, 29(6), 1102-1125.

Meyer, J. (2000). Performance with tables and graphs: Effects of training and a visual search model. *Ergonomics*, 43(11), 1840-1865.

Meyer, J., Shinar, D., & Leiser, D. (1997). Multiple factors that determine performance with tables and graphs. *Human Factors*, 39(2), 268-286.

Moore, P. J. (1993). Metacognitive processing of diagrams, maps, and graphs. *Learning and Instruction*, 3, 215-226.

Moray, N. (1969). *Listening and Attention*. Baltimore, MD: Penguin Books.

Newell, A., & Rosenbloom, P.S. (1981). Mechanisms of skill acquisition and the power law of practice. In J.R. Anderson (Ed.), *Cognitive Skills and Their Acquisition* (pp. 1-55). Hillsdale, NJ: Erlbaum.

Nyquist, H. (1928). Certain topics in telegraph transmission theory. *Transactions of AIEE*, 47, 617-644.

Oliver, F. (1998). *How to present information in graphs and diagrams* [Brochure]. London: Royal Statistical Society.

Palomaki, H. (2006). Meanings conveyed by simple auditory rhythms. *Proceedings of the 12th International Conference on Auditory Display (ICAD2006)*, London, UK.

Peden, B. F., & Hausmann, S. E. (2000). Data graphs in introductory and upper psychology textbooks: A content analysis. *Teaching of Psychology*, 27(2), 93-93.

Peres, S. C., & Lane, D. M. (2003). Sonification of statistical graphs. *Proceedings of the International Conference on Auditory Display (ICAD2003)*, Boston, MA.

Peres, S. C., & Lane, D. M. (2005). Auditory graphs: The effects of redundant dimensions and divided attention. *Proceedings of the International Conference on Auditory Display (ICAD2005)* (pp. 169-174), Limerick, Ireland.

Pinker, S. (1990). A theory of graph comprehension. In R. Freedle (Ed.), *Artificial Intelligence and the Future of Testing.* (pp. 73-126). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Pirhonen, A., Murphy, E., McAllister, g., & Yu, W. (2006). Non-speech sounds as elements of a use scenario: A semiotic perspective. *Proceedings of the 12th International Conference on Auditory Display (ICAD2006)*, London, UK.

Povel, D. J., & Essens, P. (1985). Perception of temporal patterns. *Music Perception*, 2(4), 411-440.

Quesenbery, W. (1999). Usability interface web accessibility inititative. Retrieved March 11, 2006, from http://www.stcsig.org/usability/newsletter/9910-wai.html

Resnikoff, S., Pascolini, D., Etya'ale, D., Kocur, I., Pararajasegaram, R., Pokharel, G. P., et al. (2004). Global data on visual impairment in the year 2002. *Bulletin of the World Health Organization*, 82(11), 844-851.

Roth, P., Kamel, H., Petrucci, L., & Pun, T. (2002). A comparison of three nonvisual methods for presenting scientific graphs. *Journal of Visual Impairment & Blindness*, 96(6), 420-428.

Sanderson, P. M., Flach, J. M., Buttigieg, M. A., & Casey, E. J. (1989). Object displays do not always support better integrated task performance. *Human Factors*, 31(2), 183-198.

Schutz, H. G. (1961). An evaluation of formats for graphic trend displays--Experiment II. *Human Factors*, 3, 99-107.

Shannon, C. E. (1998/1949). Communication in the presence of noise. *Proceedings of the IEEE*, 86(2), 447-457.

Smith, D. R., & Walker, B. N. (2002). Tick-marks, axes, and labels: The effects of adding context to auditory graphs. *Proceedings of the International Conference on Auditory Display (ICAD2002)* (pp. 362-367), Kyoto, Japan.

Smith, D. R., & Walker, B. N. (2005). Effects of auditory context cues and training on performance of a point estimation sonification task. *Applied Cognitive Psychology*, 19(8), 1065-1087.

Stockman, T., Hind, G., & Frauenberger, C. (2005). Interactive sonification of spreadsheets. *Proceedings of the International Conference on Auditory Display (ICAD2005)* (pp. 134-139), Limerick, Ireland.

Takagi, H., Asakawa, C., Fukuda, K., & Maeda, J. (2004). Accessibility designer: Visualizing usability for the blind. *Proceedings of the ASSETS '04* (pp. 177-184), Atlanta, GA.

Turnage, K. D., Bonebright, T. L., Buhman, D. C., & Flowers, J. H. (1996). The effects of task demands on the equivalence of visual and auditory representations of periodic numerical data. *Behavior Research Methods, Instruments & Computers*, 28(2), 270-274.

Turnbull, W. W. (1944). Pitch discrimination as a function of tonal duration. *Journal of Experimental Psychology*, 34, 302-316.

van Valkenburg, D., & Kubovy, M. (2003). In defense of the theory of indispensable attributes. *Cognition*, 87, 225-233.

W3C. (2000). Core techniques for Web Accessibility Guidelines 1.0. Retrieved March 12, 2006, from http://www.w3.org/TR/WCAG10-CORE-TECHS/#toc

Walker, B. N. (2002). Magnitude estimation of conceptual data dimensions for use in sonification. *Journal of Experimental Psychology: Applied*, 8, 211-221.

Walker, B. N., & Cothran, J. T. (2003). Sonification Sandbox: A graphical toolkit for auditory graphs. *Proceedings of the International Conference on Auditory Display (ICAD2003)* (pp. 161-163), Boston, MA.

Walker, B. N., & Lowey, M. (2004). Sonification Sandbox: A graphical toolkit for auditory graphs. *Proceedings of the Rehabilitation Engineering & Assistive Technology Society of America (RESNA) 27th International Conference*, Orlando, FL.

Walker, B. N., & Nees, M. A. (2005a). An agenda for research and development of multimodal graphs. *Proceedings of the International Conference on Auditory Display (ICAD2005)* (pp. 428-432), Limerick, Ireland.

Walker, B. N., & Nees, M. A. (2005b). Conceptual versus perceptual training for auditory graphs. *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting* (pp. 1598-1601), Orlando, FL.

Wenzel, E. M., Arruda, M., Kistler, D. J., & Wightman, F. L. (1993). Localization using nonindividualized head-related transfer functions. *Journal of the Acoustical Society of America*, 94(1), 111-123.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, 3(2), 159-177.

Zacks, J., Levy, E., Tversky, B., & Schiano, D. (2002). Graphs in print. In M. Anderson, B. Meyer & P. Olivier (Eds.), *Diagrammatic Representation and Reasoning* (pp. 187-206). London: Springer.